JOURNAL OF BACTERIOLOGY, Dec. 1995, p. 6952–6957 Vol. 177, No. 23
0021-9193/95/$04.00+0

Cloning and Characterization of the *sigA* Gene Encoding the Major Sigma Subunit of *Rhizobium meliloti*

BRENDA G. RUSHING† AND SHARON R. LONG*

Howard Hughes Medical Institute, Department of Biological Sciences, Stanford University, Stanford, California

Received 7 July 1995/Accepted 29 September 1995

Using PCR to create a probe based on conserved region 2 of sigma factors, we have cloned the *sigA* gene coding for the major sigma factor of *Rhizobium meliloti*. The 684-residue protein encoded by the *sigA* gene was expressed in vitro in coupled transcription-translation experiments with *R. meliloti* extracts and migrated aberrantly in sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Its deduced amino acid sequence is similar to that of RpoD of *Escherichia coli* and is nearly identical to that of SigA of the closely related bacterium *Agrobacterium tumefaciens*. Through Southern analysis, we located the gene on the *R. meliloti* main chromosome rather than on one of the megaplasmids. The *sigA* locus does not appear to be part of a macromolecular synthesis operon (MMS), as in many other bacterial species, but rather lies downstream of a partial open reading frame showing similarity to the threonine dehydrogenase gene (*tdh*) of *E. coli*.

Rhizobium meliloti grows as either a free-living organism or a symbiont with plants. Entry into symbiosis involves a series of developmental transitions, each of which is accompanied by characteristic changes in gene expression (22). Early in the interaction between plant and bacteria, the bacteria express nodulation (*nod*) genes under stimulation of a plant inducer, mediated by the transcriptional activator NodD. The final differentiated state of the bacterium is a nitrogen-fixing bacteroid that displays distinctive gene expression including transcription of *nif* and other genes for bacteroid metabolism and nitrogen fixation. The FixL-FixJ two-component system leads to activation of a series of transcriptional regulators including NifA and FixK, which direct expression of the symbiotic genes. At intermediate developmental stages, the bacteria invade the plant through infection threads. It is not yet known whether distinctive gene expression patterns exist for these stages as well.

The transcription activators for symbiotic gene expression have been identified by mutation, cloning, and sequence analyses and promoter binding studies. However, relatively little is known about the biochemical mechanisms of transcription control, either during symbiotic development or for housekeeping functions. Improved understanding will require study of the promoter requirements for constitutive and regulated transcription, including the properties of the RNA polymerases. The sigma (σ) factor proteins of RNA polymerase have attracted attention because of their fundamental role in promoter recognition. In previous work, we observed that *R. meliloti* RNA polymerase had promoter requirements in vitro that were different from those of the *Escherichia coli* RNA polymerase (14). In this study, we sought to characterize the housekeeping σ factor component of *R. meliloti* RNA polymerase as a step towards studying the requirements for promoter recognition. By means of a PCR approach, we generated a σ factor probe with which we isolated the *sigA* gene encoding the primary σ factor of *R. meliloti*. We have shown expression of SigA in vitro and its aberrant electrophoretic migration, consistent with behavior of other σ factors. *R. meliloti sigA* does not appear to be encoded within the highly regulated macromolecular synthesis (MMS) operon found in *E. coli* and many other bacteria and is located on the *R. meliloti* chromosome rather than on one of the symbiotic megaplasmids.

* Corresponding author. Mailing address: Howard Hughes Medical Institute, Department of Biological Sciences, Stanford University, Stanford, CA 94305-5020. Phone: (415) 723-3232. Fax: (415) 725-8309.

† Present address: Department of Biology, Indiana University, Bloomington, IN 47405.

MATERIALS AND METHODS

Strains, plasmids, and growth conditions. Bacterial strains and plasmids used in this work are listed in Table 1. Bacterial cultures were grown in Luria-Bertani or TY medium as previously described (4) supplemented with the appropriate antibiotics as follows: ampicillin, 50 μg/ml; tetracycline, 10 μg/ml; streptomycin, 500 μg/ml. *E. coli* strains were grown at 37°C; *R. meliloti* strains were grown at 30°C. *E. coli* cultures grown for λ plating were supplemented with final concentrations of 0.2% maltose and 10 mM magnesium sulfate. Platings were performed with either NZCYM or NZY top agarose (30).

DNA manipulations and sequencing. Restriction digestion, transformation, and DNA isolation were performed by standard methods (30). The majority of the nucleotide sequence was obtained by automated sequencing of conveniently sized pBGR48 subclones on an Applied Biosystems model 373A (Protein and Nucleic Acid Facility, Stanford University). The Sequenase version 2.0 DNA sequencing kit (U.S. Biochemical Corporation, Cleveland, Ohio) was used to clarify any sequence discrepancies. The locations of all restriction sites used for subcloning were confirmed by determining the sequence across those sites.

PCR. Partially degenerate oligonucleotide primers were based on conserved regions 2.1 and 2.3/2.4 of sigma factors (9), with added cloning sites at the 5′ ends. The sequences of primers 1 and 2 were 5′-GGCTCGAGAT(ACT)GC NAA(AG)(AC)GNTA-3′ and 5′-CCAAGCTTGC(CT)TGNC(GT)(AGT)ATC CACC-3′, respectively. Primer 1 has an *Xho*I restriction site; primer 2 contains a *Hin*dIII site. Reaction mixtures contained 8 μM each primer, 0.2 mM deoxynucleoside triphosphates, 1 μg of *R. meliloti* genomic DNA, 10 mM Tris (pH 8.3), 1.5 mM $MgCl_2$, 50 mM KCl, 0.1 mg of gelatin per ml, and 3 U of *Thermus aquaticus* DNA polymerase (Cetus Corp.) in a final volume of 100 μl. The mixture was layered with mineral oil and subjected to 30 cycles of 1 min at 95°C, 1 min at 55°C, and 2 min at 74°C. Since the primary reaction gave a small amount of product, the DNA was isolated by the freeze-squeeze method (33) and subjected to another round of amplification prior to cloning into pBluescript SK(+).

Screening of λ libraries and Southern blot conditions. Nitrocellulose membranes (Schleicher and Schuell) were used for plaque lifts. Filters were treated by either an autoclave lysis-fixation method in which they were wrapped loosely in blotting paper and autoclaved for 2 min on the liquid cycle or by standard methods using sequential wetting with denaturation and neutralization buffers followed by baking under vacuum at 80°C (30). Magnagraph nylon membranes (MSI) were used for Southern transfers. DNA probes were labeled by random hexamer priming (11). Membranes were hybridized at 65°C in 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–1× SSPE (0.18 M NaCl, 10 mM NaH_2PO_4, and 1 mM EDTA [pH 7.7])–5× Denhardt's solution–0.1% sodium dodecyl sulfate (SDS)–0.5 mg of denatured salmon sperm DNA per ml. Washes were done at room temperature in 1× SSC–0.25% SDS. The two λ libraries (constructed by D. Bramhill) used for this study contained *R. meliloti* 1021 genomic DNA digested with *Bam*HI ligated into LambdaGEM-11 (Promega) or *R. meliloti* DNA digested with *Sac*I ligated into LambdaGEM-12 (Promega).

Construction of the full-length *sigA* ORF. The initial λ isolate from the *Bam*HI

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source or reference
Strains		
E. coli		
XL1-Blue	*recA lac* F′ Tc^r	Stratagene
LE392	λ plating strain, *rec*$^+$	Promega
KW251	λ plating strain, *recD*	Promega
R. meliloti		
RCR2011	Wild type	25
1021	RCR2011 derivative, Str^r	26
A. tumefaciens A136	Nal^r Rif^r derivative of NT1, a Ti plasmid-cured derivative of C-58	17, 35
Plasmids		
pBluescript SK(+)	Cloning vector, Ap^r	Stratagene
pMRG8	*E. coli rpoD*	18
pJSS56	4.3-kb *Hin*dIII fragment in pUC118, contains *nodPQ*$_2$	31
pBGR8	146-bp PCR product, *Xho*I-*Hin*dIII insert in SK(+)	This study
pBGR12	2.6-kb *Bam*HI insert in SK(+), contains the 5′ end of *sigA*	This study
pBGR45b	1.2-kb *Sst*I insert in SK(+), contains the 3′ end of *sigA*, overlaps pBGR12	This study
pBGR48	3.4-kb insert in SK(+), contains full-length *sigA*[a]	This study
pBGR61a	2.5-kb *Hin*dIII insert in SK(+), contains full-length *sigA* from pBGR48	This study

[a] See the text for construction.

library contained a 2.6-kb fragment in which the putative *sigA* open reading frame (ORF) was truncated at an internal *Bam*HI site; a pBluescript SK(+) subclone was named pBGR12. Restriction analysis of this clone for *Sst*I sites and Southern analysis for hybridization of the probe to genomic *Sst*I-cut DNA predicted a 1.2-kb overlapping clone, which we obtained by screening a *Sac*I (*Sst*I) library [cloned into pBluescript SK(+) as pBGR45b]. These plasmids were digested to completion with *Eco*RI (polylinker sites) and then subjected to partial *Sst*I digestion. The 5.0-kb band from the pBGR12 digestion was isolated and subjected to further digestion with *Pst*I to remove any DNA that could religate and give a high number of background colonies. The 1.2-kb *Sst*I-*Eco*RI band from the pBGR45b digestion was isolated and inserted into the pBGR12 vector, creating pBGR48.

Expression of SigA protein in vitro. Plasmid DNA (1 μg) purified by using a Qiagen plasmid miniprep kit (Qiagen Inc.) was incubated with S-30 extract from *R. meliloti* RCR2011 as described previously (16). Protein products were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by autoradiography.

Sequence analysis. Sequence manipulations, comparisons, and database searches were performed by using the University of Wisconsin Genetics Computer Group programs (7).

Nucleotide sequence accession number. The nucleotide sequence date reported in this paper have been submitted to GenBank and assigned the accession number L47288.

RESULTS

Cloning of the *sigA* gene from *R. meliloti*. Sigma factors contain several regions which are highly conserved at the amino acid level (21). As the segment designated region 2 is the most conserved portion of σ proteins, we chose to use a PCR approach to amplify *R. meliloti*-specific sigma gene sequences using primers based on subregions 2.1 and 2.3/2.4 (9). Amplification of *R. meliloti* DNA using these primers gave rise to a 146-bp product cloned as pBGR8. The nucleotide sequence of the insert revealed 81% identity to the *rpoD* gene encoding the primary σ^{70} subunit of *E. coli* and an inferred

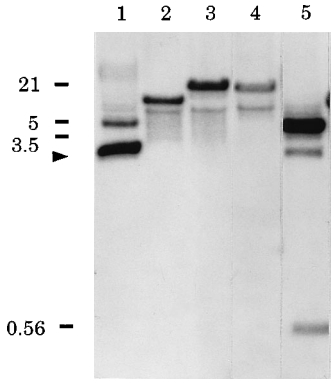

FIG. 1. Southern analysis of the *R. meliloti sigA* gene. Total DNA was digested with *Bam*HI, *Eco*RI, *Hin*dIII, *Bgl*II, or *Pst*I (lanes 1 to 5, respectively). The membrane was probed with the 146-bp insert from pBGR8. Numbers on the left indicate sizes of selected λ markers in kilobases. Arrowhead, 2.6-kb *Bam*HI fragment corresponding to the insert in pBGR12. Note that lane 5 contains three hybridizing bands; because there is a *Pst*I site within the PCR product, the pattern of three bands is consistent with recognition of two σ sequences.

partial ORF of 48 amino acids that showed 95% identity to a corresponding segment of the *E. coli* protein. We later determined that the two amino acid differences in the PCR product are not actually present in the genomic version of the *R. meliloti* gene and were changes introduced by the degeneracy of the primers. The bias of the amino acid sequence towards usage of preferred *R. meliloti* codons suggested that the amplified DNA was of *R. meliloti* origin. While a full-length *E. coli* σ^{70} probe (3-kb *Pvu*II fragment from pMRG8) hybridized only faintly to *R. meliloti* DNA and would not have been useful to screen a library (data not shown), the smaller, highly conserved PCR product did hybridize to DNAs from both bacteria. However, this cross-hybridization did not cause a background problem during the library screenings, presumably because of the much higher copy number of the desired *R. meliloti* sequence within a positive plaque.

Analysis of a Southern blot showed that the PCR product routinely hybridized to two bands of *R. meliloti* DNA with differing intensities, suggesting the presence of two different sequences related to σ factor genes (Fig. 1). We screened a λ library containing *Bam*HI-digested *R. meliloti* DNA in an attempt to obtain both a strongly hybridizing *Bam*HI fragment of 2.6 kb and a weakly hybridizing one of 4.8 kb. We isolated multiple clones containing the 2.6-kb insert and other fragments later shown to be noncontiguous. The 2.6-kb fragment was subcloned as pBGR12, and initial sequence data showed that the similarity to σ^{70} was truncated at the cloning site. On the basis of similarity to *E. coli rpoD*, it appeared that we had isolated approximately two-thirds of the 5′ end of the gene. We did not recover any λ clones containing the 4.8 fragment, because the autoclave method of preparing the filters markedly decreases the sensitivity of the screen.

From initial sequence and restriction mapping of the pBGR12 insert, we determined that an *Sst*I site lay 370 bp upstream of the *Bam*HI site and that the pBGR8 insert (i.e., the cloned PCR product) would hybridize within that region. A Southern blot indicated that the overlapping genomic *Sst*I band was about 1.2 kb in length (data not shown) and should thus be large enough to contain the remaining σ factor ORF. Therefore, we screened a λ library made from *Sac*I-digested *R. meliloti* DNA and isolated five clones containing the 1.2-kb insert. This fragment was cloned as pBGR45b and was used to

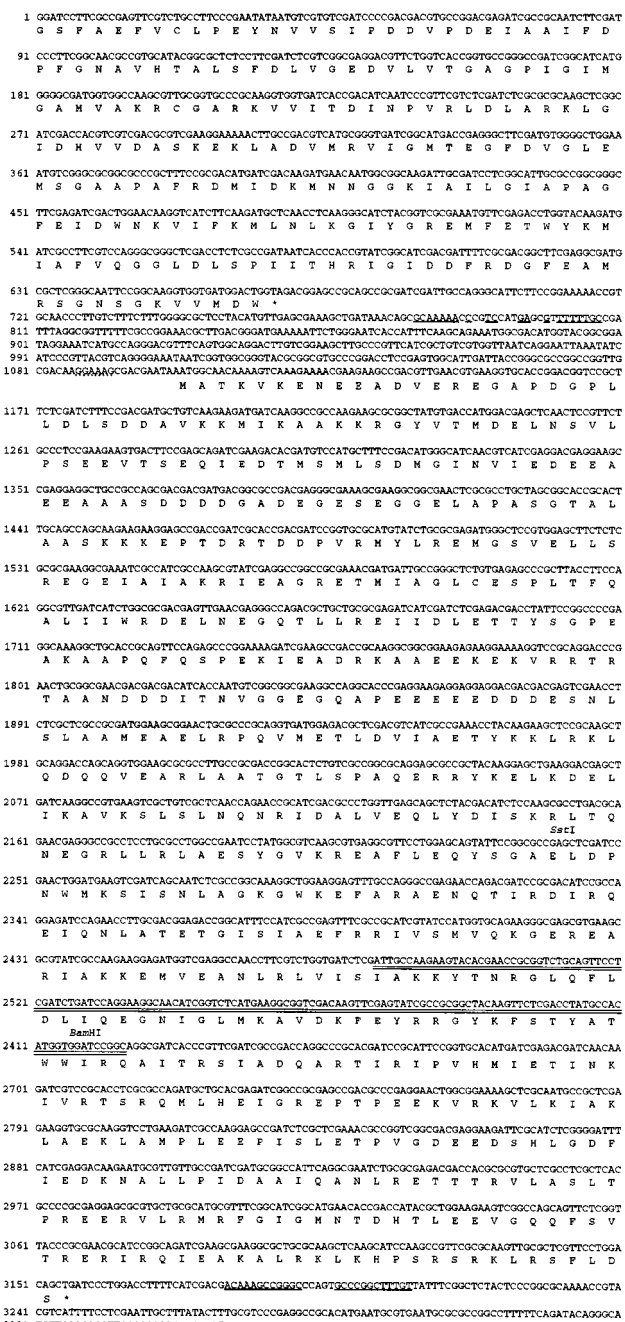

FIG. 2. Nucleotide sequence and predicted amino acid sequence of the *R. meliloti sigA* gene. The 3′ sequence of the *tdh* gene and its deduced amino acid sequence are shown upstream of *sigA*. The locations of the PCR product within *sigA*, a putative Shine-Dalgarno ribosome binding site, and possible rho-independent transcription terminator sites between *tdh* and *sigA* and also following *sigA*, as determined by the Terminator program (5), are underlined with two lines, a dotted line, and a continuous line, respectively. The *Bam*HI site that corresponds to the end of pBGR12 and the *Sst*I site used to create the full-length *sigA* are indicated.

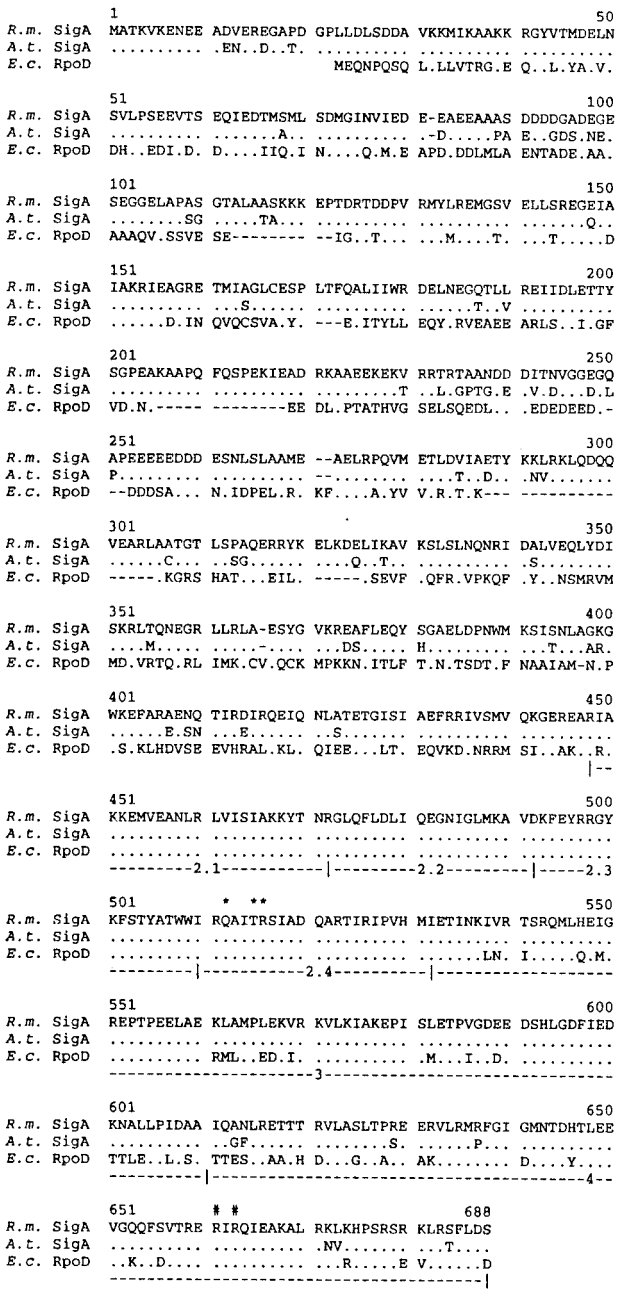

FIG. 3. Comparison of amino acid sequences of primary σ factors from *R. meliloti* (R.m.), *A. tumefaciens* (A.t.), and *E. coli* (E.c.). The entire *R. meliloti* SigA sequence is presented on the first line, and only the divergent residues are shown for the other proteins. Residues identical to those in *R. meliloti* (dots) and gaps introduced by the Pileup program (University of Wisconsin Genetics Computer Group) (dashes) are indicated. Regions 2 through 4 (21) are indicated below the sequences. Residues known to contact the −10 (∗) and −35 (#) promoter regions in other bacteria are also indicated.

create a complete *sigA* gene in pBGR48, as described in Materials and Methods.

Nucleotide sequence and predicted amino acid sequence of *sigA*. We determined the DNA sequence of the 3.4-kb insert of pBGR48 (Fig. 2) and verified the location of an ORF encoding a 684-amino-acid protein with a predicted size of 77 kDa. This ORF uses a high number of codons preferred by *R. meliloti* (Codonpreference; University of Wisconsin Genetics Computer Group) and has a possible ribosome binding site upstream (Fig. 2), indicating that it is likely to be well expressed in vivo. A putative rho-independent terminator lies 22 bp downstream of the stop codon. Since the protein encoded by the ORF is 90% identical to SigA from *Agrobacterium tumefaciens* and 51% identical to RpoD from *E. coli*, this *sigA* gene

appears likely to code for the primary σ subunit of *R. meliloti*. A comparison of these three proteins is shown in Fig. 3. The *R. meliloti* and *A. tumefaciens* proteins are 70 amino acids longer than the *E. coli* σ^{70}. Most of the loss of similarity is in the amino-terminal region 1, which is the least conserved portion of σ proteins. In regions 2 to 4, the *R. meliloti* protein is 97% identical to that of *A. tumefaciens* and 83% identical to the *E. coli* protein. We confirmed the *sigA* sequence as that encoding the *R. meliloti* primary σ factor by direct comparison of the inferred ORF with the amino-terminal sequence of the protein corresponding to the σ subunit (13) of an active preparation of purified *R. meliloti* RNA polymerase (12). The peptide sequence, TKVKEN, corresponds to amino acids 3 through 8 encoded by the *sigA* gene (12).

The *rpoD* gene in many enteric bacteria lies in an operon with the genes encoding ribosomal protein S21 (*rpsU*) and DNA primase (*dnaG*) (23, 34). Because the products of these genes are involved in formation of major cell constituents, this operon has been termed MMS. We searched the sequence surrounding the *R. meliloti sigA* gene and failed to find homology to other MMS operon genes. Instead, we found the 3′ end of an ORF ending 437 bp upstream of *sigA* that was significantly similar to the threonine dehydrogenase (*tdh*) gene of *E. coli* (Fig. 2). In contrast, no ORFs were found upstream of *sigA* in *A. tumefaciens* (32). The partial *tdh* ORF uses a high number of preferred *R. meliloti* codons and is 70% identical at the amino acid level to the carboxy-terminal two-thirds of the *E. coli* protein. Because of the distance between these two genes, it appears unlikely that they are expressed as an operon. A search for a rho-independent terminator in the intergenic sequence revealed a possible termination stem-loop structure 110 bp downstream of the *tdh* stop codon (Fig. 2).

Expression of the SigA protein in vitro. We confirmed that the spliced pBGR48 contained all of *sigA* by expressing the intact *sigA* gene product in a *R. meliloti* coupled transcription-translation system. As shown in Fig. 4, a protein product doublet that migrates at an apparent size of 98 kDa is produced from two constructs containing the full-length gene. The doublet bands are most likely an artifact of the S-30 system and are sometimes seen in assays of other genes. The size of this product corresponds to that of the σ subunit from a functional, purified *R. meliloti* RNA polymerase (12, 13). The plasmid containing *sigA* truncated at the internal *Bam*HI site produces a protein doublet that migrates at approximately 82 kDa. This construct codes for the 506 C-terminal amino acids of SigA (74% of the ORF) plus 82 amino acids from the vector sequence for a total fusion protein of 588 residues with an expected size of approximately 66 kDa. These in vitro-produced SigA proteins migrate on an SDS-polyacrylamide gel at sizes much larger than expected from the deduced amino acid sequence. This observation is consistent with the previously described behavior of other σ subunits (19). The pBGR61a template produces the SigA protein in the S-30 system, which either may arise from an endogenous *R. meliloti* promoter located within the 160 nucleotides upstream of the *sigA* start codon or may reflect expression from the *lac* promoter on the vector. The transcription start site of *A. tumefaciens sigA* has been determined (32). Comparison of DNA around the mapped *A. tumefaciens* promoter with the sequence upstream of *R. meliloti sigA* shows a nearly identical sequence through the putative Shine-Dalgarno site but little similarity further upstream; thus, no putative promoter sequences can be identified by inspection.

Hybridization analysis to determine the replicon location of the *sigA* gene. *R. meliloti* contains two large plasmids called pSyma and pSymb in addition to the chromosome. These

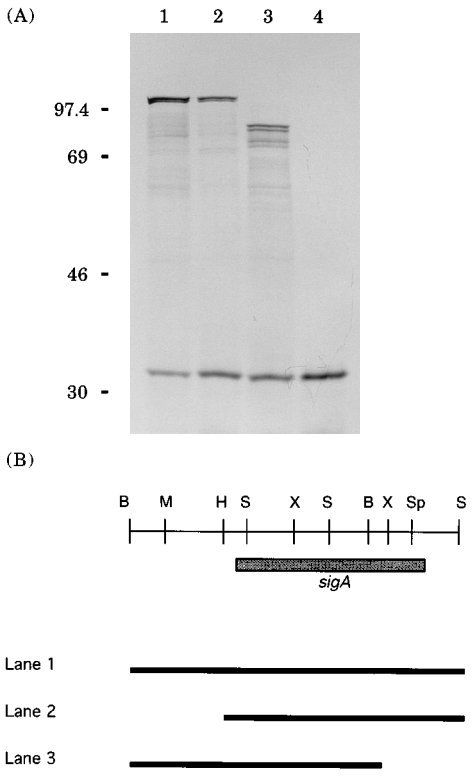

FIG. 4. Expression of SigA in vitro. (A) SDS-PAGE of coupled transcription-translation assays. The sizes (in kilodaltons) of molecular weight markers (Rainbow markers; Amersham) are indicated on the left. Lane 1, pBGR48; lane 2, pBGR61a; lane 3, pBGR12; lane 4, vector control. (B) Restriction map of the *sigA* region and schematic diagram of constructs used in protein expression analysis. Shaded box, location of the *sigA* ORF within the fragment. The DNAs included in the constructs are represented by horizontal lines below. Restriction sites: B, *Bam*HI; M, *Mlu*I; H, *Hin*dIII; S, *Sst*I; X, *Xho*I; Sp, *Sph*I.

megaplasmids are about 1,400 and 1,600 kb, respectively (3, 6, 20). Strains of *A. tumefaciens* containing each of the *R. meliloti* Sym plasmids separately can be used as sources of DNA for a Southern blot to permit rapid determination of the replicon location of a gene (31). We probed DNAs from these strains with a full-length *R. meliloti sigA* probe (Fig. 5). In lane 1, this probe hybridizes to a 10-kb *Eco*RI fragment of *R. meliloti* DNA. Lanes 2 to 4 show that only the *Eco*RI bands corresponding to the *A. tumefaciens* version of *sigA* are detected in strains containing the megaplasmids. Thus, it appears that *sigA* is encoded on the chromosome of *R. meliloti* and not on one of the Sym plasmids. Since this conclusion is based on the lack of a hybridizing fragment, we confirmed the presence of pSyma and pSymb DNA in lanes 2 and 3 by probing the same blot with a probe for $nodP_2$ (Fig. 5B). This probe detected the expected bands for both $nodP_1$ on pSyma (lane 2) and $nodP_2$ on pSymb (lane 3) compared with the *R. meliloti* total DNA (lane 1). A more faintly hybridizing *A. tumefaciens*-specific band is also seen in lanes 2 to 4, suggesting a possible $nodP_2$ homolog or the locus for ATP sulfurylase in *A. tumefaciens.*

DISCUSSION

The PCR proved a useful tool to create an *R. meliloti*-derived probe similar to conserved region 2 of σ subunits of RNA polymerase. We isolated and performed initial characterization of the gene encoding the vegetative σ factor, *sigA*, from *R. meliloti*. The deduced amino acid sequence of the SigA

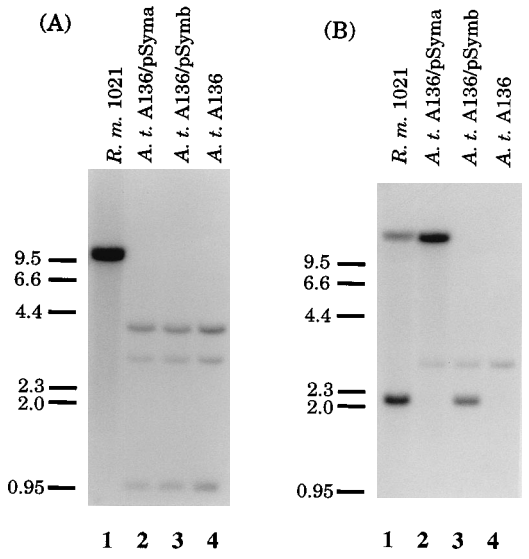

FIG. 5. Southern analysis to determine replicon location of the *sigA* gene. R.m., *R. meliloti*; A.t., *A. tumefaciens*. DNA samples were digested with *Eco*RI. The sizes (in kilobases) of selected λ markers are indicated. The same membrane was hybridized separately with the indicated probes. (A) *R. meliloti sigA* (2.5-kb *Hin*dIII fragment from pBGR48); (B) *R. meliloti* $nodP_2$ (320-bp *Eco*RI-*Sst*I fragment from pJSS56).

protein is very similar to those of *E. coli* RpoD and other members of the σ^{70} family (21). In particular, the 684-amino-acid *R. meliloti* protein is nearly identical to σ^{77} from the closely related species *A. tumefaciens* (32). Our *sigA* gene is expressed in an in vitro *R. meliloti* coupled transcription-translation system and, as is typical for σ proteins, migrates more slowly than expected on an SDS-polyacrylamide gel. To confirm that the gene we have isolated codes for the functional σ factor in RNA polymerase, we showed correspondence of residues 3 to 8 of the deduced SigA protein sequence with the amino-terminal peptide sequence of σ factor from a purified *R. meliloti* holoenzyme (12). Genomic hybridization patterns and further studies (29a) indicate the existence of other σ factor sequences in *R. meliloti*. Both primary sequence data and the size of the encoded product indicate that the other σ sequences are distinct from *sigA* (29).

The σ factor is a critical component of the RNA polymerase holoenzyme that allows the complex to distinguish a promoter sequence within the DNA. Mutational analysis of sigmas and the promoters they recognize has shown that residues in the defined region 2.4 interact at the −10 of a promoter while residues in region 4.2 are involved in −35 interaction (for a review, see reference 21). However, other determinants must also be at work, considering the conservation in these regions among bacteria for which it is known that heterologous promoters are not recognized in vivo. The *A. tumefaciens* and *R. meliloti* SigA proteins and the *E. coli* RpoD are all identical at the known recognition residues, specifically Q-508, T-511, and R-512 in region 2.4 and R-655 and R-659 in region 4.2 (Fig. 3; *R. meliloti* numbering).

The biochemical and genetic characterization of σ factors will make it possible to explore in more detail the interactions of RNA polymerase with promoters, including those of both housekeeping genes and developmentally regulated genes. Relatively few transcription start sites have been mapped in *R. meliloti*. Since this bacterium has been studied mainly for its symbiotic nodulation (*nod*) and nitrogen-fixing (*nif*) abilities with the host plant *Medicago sativa* (alfalfa), many of the genes investigated direct very specialized functions for the cell. Thus, a typical "housekeeping" consensus promoter has not been defined for *R. meliloti*, although several genes presumed to be transcribed by a housekeeping RNA polymerase holoenzyme (*hemA*, *nodD*, *ntrA*, and *trpE*) have a conserved TTG sequence in the −35 promoter region that is also seen in *E. coli* σ^{70}-type promoters. The −10 regions of these promoters show no consensus. The *fla* genes of *R. meliloti* have a promoter region that resembles that of the *E. coli* and *B. subtilis* flagellar genes (27) and thus may require a σ^{28} homolog (FliA or SigD). The *nif* genes have a consensus −12/−24 promoter that requires σ^{54} (NtrA or RpoN) to direct transcription (28).

Different transcription requirements for *E. coli* and *R. meliloti* RNA polymerases are indicated by studies of the *trp* and *nodD1* promoters. The *R. meliloti trp* promoter corresponds only partially (50% identical positions) to the *E. coli* consensus (1) and is not expressed in *E. coli*. However, *trp* promoter mutations in the −10 region that approach the *E. coli* consensus allow *trp* to be expressed in *E. coli* (2). In contrast, the *E. coli trp* promoter directs in vitro transcription by *R. meliloti* RNA polymerase (13). In the case of nodD, it has been shown that the constitutive *nodD1* promoter functions in *A. tumefaciens* (36). However, this *nodD1* gene is not expressed by *E. coli* RNA polymerase either in vivo (36) or in vitro (14), even though the sigma factor residues known to contact promoter sequences in σ^{77} of *R. meliloti* and σ^{70} of *E. coli* are identical. Conservation at these residues has also been reported for SigA of *Chlamydia trachomatis*, although the promoters described for this pathogen also diverge from those of *E. coli* and do not function in that bacterium (9). Thus, further examination of housekeeping promoter requirements and of the relationship of σ structure to promoter recognition is needed.

The mechanism of expression for inducible *nod* genes is also not well understood. The promoters have been found to include a highly conserved sequence, known as the *nod* box, which is bound by a family of positive transcriptional activators, the NodD proteins (15). The *nod* box overlaps the −35 region of these promoters and, as such, is highly conserved. However, the −10 regions of these promoters do not share a consensus sequence (14). These promoters also do not mimic the standard *E. coli* promoter. It has not been determined which σ subunit is involved in the transcription of the inducible *nod* genes. The σ^{77} protein is an attractive candidate for two reasons: first, it is shared by *R. meliloti* and *A. tumefaciens*, both of which can support expression of inducible *nod* gene promoters (36), and second, the σ^{77} protein is present in a preparation of *R. meliloti* RNA polymerase that can initiate transcription of the *nod* genes (15a).

Cloning of the *sigA* sequence allowed us to determine the location and arrangement of the gene. We found that *sigA* is located on the chromosome, as has also been shown for *ntrA* (28), which prior to this study was the only σ subunit gene to have been cloned from *R. meliloti*. Unlike genes encoding the primary σ factors from many other bacteria (34), the *R. meliloti* homolog is not found in an MMS operon with genes for a ribosomal protein (*rpsU*) and DNA primase (*dnaG*). The MMS operon is under complex regulation, reflecting its importance in the production of the cell's major informational molecules for translation, replication, and transcription (23). This organization appears to be conserved across many gram-negative enteric bacteria as well as in the nonenteric *Haemophilus influenzae* (34) and the obligate intracellular bacterium *Rickettsia prowazekii* (24). In the gram-positive *Bacillus subtilis*, the *dnaE* primase gene is upstream of *rpoD*, but a protein of unknown function (P23) is encoded in the *rpsU*

position (8). Upstream of *sigA* in *R. meliloti*, however, we identified the 3′ end of an ORF that is very similar to the threonine dehydrogenase (*tdh*) gene of *E. coli*. We did not find evidence for any ORFs beginning in the 260 nucleotides downstream of *sigA*. The putative transcription termination signal found there also suggests that *sigA* is not in an operon with downstream genes.

The gene arrangement found for *sigA* suggests that *R. meliloti* does not require coordination of its transcription, translation, and replication machinery in the same manner as that seen in other bacteria. If so, perhaps this arises because of uncoupling of transcription activity from cell division during the late stages of symbiosis (22). *R. meliloti* would thus not need the same mechanism for driving all its macromolecular functions in concert. Alternatively, coordinate control of *rpsU*, *dnaG*, and *sigA* genes may occur as a regulon coordinated in *trans.*

Another consideration is whether the presence of the upstream *tdh* gene has regulatory significance. Because of the distance between these two genes (>400 bp) and a possible transcription termination site between them (Fig. 2), it seems unlikely that these coding sequences form an operon. However, if these genes are coexpressed, the significance is not clear. In *E. coli*, the repression of *tdh* by Lrp, the leucine-responsive regulatory protein, is relieved under nitrogen-rich conditions (10), but it is unknown whether any comparable system exists in *R. meliloti*. Further detailed study of *sigA* gene expression should reveal whether it is coordinately transcribed with upstream genes, and, if so, what cellular conditions influence this expression.

ACKNOWLEDGMENTS

We thank all members of our laboratory group for helpful discussions. Specifically, we thank Melanie Barnett for construction of the Codonpreference table for *R. meliloti*, Michael Willits for providing the DNA for the experiment shown in Fig. 5, David Bramhill for the λ library, and Robert Fisher for the amino-terminal sequence data of σ^{77}. We thank Robert Fisher for reading and criticizing the manuscript and JoAnne Connelly for preparing and proofing it.

B.G.R. was supported by an NIH training grant in cell and molecular biology to Stanford University. This work was also supported by NIH grant GM30962 to S.R.L.

REFERENCES

1. **Bae, Y. M., E. Holmgren, and I. P. Crawford.** 1989. *Rhizobium meliloti* anthranilate synthase gene: cloning, sequence and expression in *Escherichia coli*. J. Bacteriol. **171:**3471–3478.
2. **Bae, Y. M., and G. V. Stauffer.** 1991. Mutations that affect activity of the *Rhizobium meliloti trpE*(*G*) promoter in *Rhizobium meliloti* and *Escherichia coli*. J. Bacteriol. **173:**5831–5836.
3. **Banfalvi, Z., E. Kondorosi, and A. Kondorosi.** 1985. *Rhizobium meliloti* carries two megaplasmids. Plasmid **13:**129–138.
4. **Beringer, J. E.** 1974. R-factor transfer in *Rhizobium leguminosarum*. J. Gen. Microbiol. **84:**188–198.
5. **Brendel, V., and E. N. Trifonov.** 1984. A computer algorithm for testing potential prokaryotic promoters. Nucleic Acids Res. **12:**4411–4427.
6. **Charles, T. C., and T. M. Finan.** 1991. Analysis of a 1600-kilobase *Rhizobium meliloti* megaplasmid using defined deletions generated *in vivo*. Genetics **127:**5–20.
7. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
8. **Doi, R. H., and L.-F. Wang.** 1986. Multiple procaryotic ribonucleic acid polymerase sigma factors. Microbiol. Rev. **50:**227–243.
9. **Engel, J. N., and D. Ganem.** 1990. A polymerase chain reaction-based approach to cloning sigma factors from eubacteria and its application to the isolation of a sigma-70 homolog from *Chlamydia trachomatis*. J. Bacteriol. **172:**2447–2455.
10. **Ernsting, B. R., M. R. Atkinson, A. J. Ninfa, and R. G. Matthews.** 1992. Characterization of the regulon controlled by the leucine-responsive regulatory protein in *Escherichia coli*. J. Bacteriol. **174:**1109–1118.
11. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:**6–13.
12. **Fisher, R. F.** Personal communication.
13. **Fisher, R. F., H. L. Brierley, J. T. Mulligan, and S. R. Long.** 1987. Transcription of *Rhizobium meliloti* nodulation genes: identification of a *nodD* transcription initiation site *in vitro* and *in vivo*. J. Biol. Chem. **262:**6849–6855.
14. **Fisher, R. F., T. T. Egelhoff, J. T. Mulligan, and S. R. Long.** 1988. Specific binding of proteins from *Rhizobium meliloti* cell-free extracts containing NodD to DNA sequences upstream of inducible nodulation genes. Genes Dev. **2:**282–293.
15. **Fisher, R. F., and S. R. Long.** 1989. DNA footprint analysis of the transcriptional activator proteins NodD1 and NodD3 on inducible *nod* gene promoters. J. Bacteriol. **171:**5492–5502.

15a. **Fisher, R. F., and S. R. Long.** Unpublished data.

16. **Fisher, R. F., J. Swanson, J. T. Mulligan, and S. R. Long.** 1987. Extended region of nodulation genes in *Rhizobium meliloti* 1021. II. Nucleotide sequence, transcription start sites, and protein products. Genetics **117:**191–201.
17. **Garfinkel, D. J., R. B. Simpson, L. W. Ream, F. F. White, M. P. Gordon, and E. W. Nester.** 1981. Genetic analysis of crown gall: fine structure map of the T-DNA by site directed mutagenesis. Cell **27:**143–153.
18. **Gribskov, M., and R. R. Burgess.** 1983. Overexpression and purification of the sigma subunit of *Escherichia coli* RNA polymerase. Gene **26:**109–118.
19. **Helmann, J. D., and M. J. Chamberlin.** 1988. Structure and function of bacterial sigma factors. Annu. Rev. Biochem. **57:**839–872.
20. **Honeycutt, R. J., M. McClelland, and B. W. S. Sobral.** 1993. Physical map of the genome of *Rhizobium meliloti* 1021. J. Bacteriol. **175:**6945–6952.
21. **Lonetto, M., M. Gribskov, and C. A. Gross.** 1992. The σ^{70} family: sequence conservation and evolutionary relationships. J. Bacteriol. **174:**3843–3849.
22. **Long, S. R.** 1989. *Rhizobium* genetics. Annu. Rev. Genet. **23:**483–506.
23. **Lupski, J. R., and G. N. Godson.** 1984. The *rpsU-dnaG-rpoD* macromolecular synthesis operon of *E. coli*. Cell **39:**251–252.
24. **Marks, G. L., and D. O. Wood.** 1993. Characterization of the gene coding for the *Rickettsia prowazekii* DNA primase analogue. Gene **123:**121–125.
25. **Meade, H. M., S. R. Long, G. B. Ruvkun, S. E. Brown, and F. M. Ausubel.** 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn*5* mutagenesis. J. Bacteriol. **149:**114–122.
26. **Ogawa, J., and S. R. Long.** 1995. The *Rhizobium meliloti groELc* locus is required for regulation of early *nod* genes by the transcription activator NodD. Genes Dev. **9:**714–729.
27. **Pleier, E., and R. Schmitt.** 1989. Identification and sequence analysis of two related flagellin genes in *Rhizobium meliloti*. J. Bacteriol. **171:**1467–1475.
28. **Ronson, C. W., B. T. Nixon, L. M. Albright, and F. M. Ausubel.** 1987. *Rhizobium meliloti ntrA* (*rpoN*) gene is required for diverse metabolic functions. J. Bacteriol. **169:**2424–2431.
29. **Rushing, B. G.** 1995. Transcription factors in *Rhizobium meliloti*: characterization of the positive regulator NodD3 and two sigma subunits, SigA and SigB. Ph.D. dissertation. Stanford University, Stanford, Calif.

29a. **Rushing, B. G.** Unpublished data.

30. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
31. **Schwedock, J. S., and S. R. Long.** 1992. *Rhizobium meliloti* genes involved in sulfate activation: the two copies of *nodPQ* and a new locus, *saa*. Genetics **132:**899–909.
32. **Segal, G., and E. Z. Ron.** 1993. Cloning, sequencing and transcriptional analysis of the gene coding for the vegetative sigma factor of *Agrobacterium tumefaciens*. J. Bacteriol. **175:**3026–3030.
33. **Tautz, D., and M. Renz.** 1983. An optimized freeze-squeeze method for the recovery of DNA fragments from agarose gels. Anal. Biochem. **132:**14–19.
34. **Versalovic, J., T. Koeuth, R. Britton, K. Gesvain, and J. R. Lupski.** 1993. Conservation and evolution of the *rpsU-dnaG-rpoD* macromolecular synthesis operon in bacteria. Mol. Microbiol. **8:**343–355.
35. **Watson, R. J., Y.-K. Chan, R. Wheatcroft, A.-F. Yang, and S. Han.** 1988. *Rhizobium meliloti* genes required for C_4-dicarboxylate transport and symbiotic nitrogen fixation are located on a megaplasmid. J. Bacteriol. **170:**927–934.
36. **Yelton, M. M., J. T. Mulligan, and S. R. Long.** 1987. Expression of *Rhizobium meliloti nod* genes in *Rhizobium* and *Agrobacterium* backgrounds. J. Bacteriol. **169:**3094–3098.