Frankia sequences exhibiting RNA polymerase promoter activity

Joyce V. Bock, Ty Battershell, James Wiggington, Theodore R. John and Jerry D. Johnson

Author for correspondence: Jerry D. Johnson. Tel: +1 307 766 3300. Fax: +1 307 766 5098.
e-mail: jdjohn@uwyo.edu

Frankia are Gram-positive, filamentous bacteria capable of fixing atmospheric dinitrogen either in the free-living state or in symbiosis with a variety of woody plants. Only a few Frankia genes have been sequenced and gene expression is not well characterized. To isolate a segment of Frankia DNA that functions as an RNA polymerase promoter, fragments of Frankia strain ArI5 genomic DNA were cloned upstream of a promoterless, Vibrio harveyi luxAB cassette. Constructs were screened for luminescence in E. coli and positive clones assayed for in vitro transcription activity with a partially purified Frankia RNA polymerase extract. Primer extension analysis of in vitro transcripts produced from one clone, GLO7, identified two major transcription start sites, TSP-1 and TSP-2, 52 bp apart. Deletion analysis then localized sequences essential for promoter activity. The upstream promoter region, GLO7p1, contains sequences resembling the N35 element of a Streptomyces promoter and the N35 and N10 elements of the canonical E. coli promoter. Also within this region are two pentamers identical to sequences near the 5′ end of the Frankia strain Cpi1 glutamine synthetase gene. The second promoter, GLO7p2, contains a putative NtrC binding site at N145 and a possible σN-RNA polymerase recognition sequence at N14 suggesting that GLO7p2 may be a nitrogen-regulated promoter. An in vivo transcript representing an ORF of 498 aa starting 64 bp downstream of the distal transcription start, TSP-1, was detected by RT-PCR. This supports the conclusion that this DNA fragment has promoter activity in vivo as well as in vitro.

Keywords: Frankia, promoter sequences, in vitro transcription

INTRODUCTION

Frankia are slow-growing, Gram-positive, filamentous bacteria of the order Actinomycetales and family Frankiaceae. They are capable of fixing atmospheric dinitrogen either in the free-living state or in symbiosis with host plants representing over 194 species distributed among 24 genera (Nazaret et al., 1991). Actinorhizal host plants are trees and woody shrubs that are distributed worldwide, but are predominantly found in temperate areas (Baker & Selig, 1984) and at higher elevations of the tropics (Tjepkema et al., 1986). They are generally pioneer species, meaning that they are the first plants to colonize ecologically disturbed or nitrogen-poor sites (Baker & Selig, 1984; Tjepkema et al., 1986). The annual nitrogen fixation rates for actinorhizal plants range from 2 to 362 kg nitrogen per hectare (Stowers, 1987) and they contribute significant amounts of reduced nitrogen to the ecosystems which they occupy. Therefore, an increased understanding of Frankia biology is of interest because of their ecological importance as well as their long term potential for extending symbiosis to novel host plants.

Comparatively little is known of Frankia molecular biology and genetics. Only a few Frankia genes that encode proteins have been cloned and completely sequenced. These include genes from the nif operon of several strains (Harriott et al., 1995; Nalin et al., 1995; Normand et al., 1988), ferredoxin subunits from strain EUIK1 (Yoo et al., 1999) and glnA (Hosted et al.,...
1993) plus glnII (Rochefort & Benson, 1990) from strain Cpi1. The tRNA operon of strain ORS020606 (Normand et al., 1992) and partial 16S tRNA sequences have been characterized from many strains (Nazar et al., 1991; Normand et al., 1996). The control of gene expression has not previously been addressed in Frankia although Cournoyer & Normand (1994) examined promoter activity of Frankia DNA in Streptomyces and Escherichia coli. Also, numerous attempts to develop a transformation system for Frankia have failed perhaps, in part, due to a lack of basic information about gene expression.

The transcriptional regulation of genes involved in nitrogen fixation (nif) and other nitrogen-regulated processes (ntr) has been investigated in rhizobia and enteric bacteria. Both sets of genes are transcribed by an RNA polymerase holoenzyme containing a sigma factor, \( \sigma^N \), encoded by the ntrA gene (Gussin et al., 1986). The recognition sequence for this holoenzyme is distinct from that of the standard E. coli promoter and is referred to here as the nif/ntr promoter. Maximal gene expression is also dependent upon sequences far upstream of the transcriptional start site (Alvarez-Morales et al., 1986; Buck et al., 1986; Reitzer & Magasanik, 1986) that act as prokaryotic enhancer elements (Hunt & Magasanik, 1985; Wedel et al., 1990). Enhancer-binding proteins are different for the two sets of genes but similar in primary structure and function. The transcriptional activator for the ntr genes is encoded by the ntrC gene and the nif activator is the product of the nifA gene (Ditta, 1989). The results presented herein suggest that Frankia may use mechanisms similar to those of the enteric bacteria and rhizobia in the expression of nitrogen-regulated genes. We have identified a cloned Frankia DNA that directs initiation of transcription in a Frankia cell-free extract from two promoter regions. The upstream promoter, TSP-1, contains sequences similar to the canonical E. coli promoter and a Streptomyces promoter element. The downstream promoter, TSP-2, is tentatively identified as an ntr promoter by the presence of sequences closely related to the consensus NtrC-binding site and the nif/ntr promoter. Using RT-PCR, a mRNA was identified in extracts from Frankia ArI5 that represents a large ORF just downstream of the TSP-1 and 2 promoter elements. This corroborates the biological significance of the in vitro studies.

**METHODS**

**Bacterial cultures.** Frankia strains were cultured in a complete broth containing 5 g yeast extract l\(^{-1}\), 100 g dextrose l\(^{-1}\) and 5 g Casamino acids l\(^{-1}\) in a mineral salt solution supplemented with 1 ml l\(^{-1}\) each of micronutrient and vitamin stock solutions (Tjepkema et al., 1980). Fresh cultures were inoculated with 1/3–1/10 vol. homogenized mycelia from a stationary phase culture. The bacteria were incubated at 28°C until a dense mycelial mat formed at the bottom of the vessel, usually 14–21 d. To disrupt mycelia and improve growth, some cultures were vigorously agitated by hand once a day. E. coli cultures were grown at 37°C in LB broth containing the appropriate antibiotic (Sambrook et al., 1989). Plasmid constructs were transformed into E. coli cells made competent for DNA uptake by CaCl\(_2\) treatment (Sambrook et al., 1989).

**Transcription extracts and assays.** Transcription extracts were prepared essentially according to the procedure of Westpheling et al. (1985). All manipulations were done at 0–4°C. Two grams (wet wt) of Frankia strain Cpi1 cells were collected by centrifugation at 6000 g for 10 min and washed twice in 80 ml aliquots of buffer A (0.01 M Tris/ HCl, pH 8.0, 1 mM EDTA, 0.3 mM DTT, 0.3 mg PMSF ml\(^{-1}\), 10%, v/v, glycerol) containing 0.1 M KCl, then resuspended in 40 ml of the same buffer. The cells were subjected to two passages through a French pressure cell at 30–40 kPa. The lysate was centrifuged at 100000 g for 1 h to remove cell debris. The supernatant was collected and loaded onto a 2 ml column of heparin-agarose at a flow rate of 10–15 ml h\(^{-1}\). The column was washed with 80 ml buffer A containing 0.1 M KCl and the bound material was eluted by raising the KCl concentration to 0.5 M. Fractions of 1 ml were collected and aliquots were assayed for protein concentration and RNA polymerase activity. Active fractions were divided into aliquots and stored at -80°C.

Transcription reactions were performed in a volume of 50 μl containing 0.5 or 1.0 μg plasmid DNA, 20 μg extract protein and 1.0 μM [\( ^{32}P \)]UTP (7.4 Tbq mmol\(^{-1}\)) in transcription buffer (40 mM Tris/HCl, pH 8.0, 10 mM MgCl\(_2\), 10 mM DTT, 50 μg BSA ml\(^{-1}\), 150 μM each of ATP, CTP and GTP). The final KCl concentration was adjusted to 200 mM for the Frankia extract, an amount that had been shown to produce maximal transcription activity (data not shown). The reactions were incubated at 37°C for 10 or 15 min then placed on ice and quenched with an equal volume of a stop solution (100 mM sodium acetate, pH 5.2, 0.4% SDS, 1 mg yeast tRNA ml\(^{-1}\)). Incorporation of \( ^{32}P \) into TCA precipitable material was determined by Cerenkov counting.

**Construction of luciferase promoter-probe clones.** Plasmid pFIT001 containing the Vibrio harveyi luxAB genes on a 4.0 kbp insert (Legocki et al., 1986) was the generous gift of Dr R. P. Legocki (Boyce Thompson Institute, Cornell University, USA). A 0.6 kbp SalI fragment (Baldwin et al., 1984) was deleted to form AP, the promoterless luciferase cassette. Total DNA from Frankia strain ArI5 was partially digested with SalI and ligated into SalI-digested AP.

**Primer extension analysis of RNA.** RNA templates for primer extension were produced according to the procedure for in vitro transcription except that unlabelled UTP was used at 150 μM and 1 U RNase Block II (Stratagene) or 20 U RNASin (Promega) RNase inhibitor was included. RNA templates were also produced from the same plasmid DNAs using 1 U E. coli RNA polymerase (US Biochemical) and reaction conditions as recommended by the supplier. Transcription reactions were incubated at 37°C for 15 min then placed on ice and quenched with an equal volume of 3 M sodium acetate (pH 4.6) containing 1 mg yeast tRNA ml\(^{-1}\). Duplicate reactions were sometimes performed and combined at this stage to increase the amount of RNA available for primer extension. Reaction mixtures were extracted sequentially with phenol, phenol/chloroform/isooamyl alcohol (50:49:1) and chloroform prior to the addition of a two- to tenfold molar excess of \( ^{32}P \)-labelled oligonucleotide primer (90–180 Bq mmol\(^{-1}\)). Primers used were either AP, 5’-CGTACGGCTT-GCAAAC-3’ or T3 (Stratagene) for the GLO and SKG7 transcripts, respectively. The template/primer mixture was precipitated with ethanol, resuspended in 30 μl 40 mM PIPES (pH 6.4), 1 mM EDTA, 0.4 M NaCl, 80% formamide, heated to 70°C for 10 min and annealed for 18–24 h at 30°C.
acids were precipitated by the addition of 170 μl H2O and 400 μl absolute ethanol and incubation at 0 °C for 1 h. Pellets were collected by centrifugation and gently resuspended in 20 μl 50 mM Tris/HCl (pH 8.3), 75 mM KCl, 10 mM MgCl2, 2 mM each dATP, dGTP, dCTP and dTTP, 10 mM DTT, 20 U RNase. Primers were extended with 200 U MMLV reverse transcriptase (Bethesda Research Labs) at 37 °C for 2 h. Un-anealled RNA was removed by addition of 1 μl 0.5 M EDTA and 1 μl 5 mM DNase-free RNase A ml⁻¹ followed by incubation at 37 °C for 30 min. The mixture was extracted with an equal volume of phenol/chloroform/isoamyl alcohol and nucleic acids precipitated with 2 vols chloroform. Pellets were resuspended in 3 μl 80% formamide, 10 mM EDTA (pH 8/0), 1 mg xylene cyanol ml⁻¹, 1 mg bromophenol blue ml⁻¹ and heated at 95 °C for 5 min before electrophoresis on a 7% polyacrylamide gel containing 8 M urea (Sambrook et al., 1989).

General molecular biology techniques. Restriction enzyme digestions were performed according to suppliers’ directions. Single-stranded oligodeoxynucleotides were labelled using T4 polynucleotide kinase and [α-³²P]ATP (Sambrook et al., 1989). Unincorporated ATP was removed using DuPont Nensorb columns according to the manufacturer’s directions. Plasmids to be used in restriction analysis, subcloning and sequencing were prepared by the alkaline lysis method and further purified by polyethylene glycol precipitation (Sambrook et al., 1989).

In the preparation of templates for transcription, RNase digestion was avoided and contaminating RNA was removed by either Sephacryl S-1000 chromatography (Raymond et al., 1988) or equilibrium centrifugation in CsCl containing ethidium bromide (Sambrook et al., 1989). DNA sequencing was performed with Tag DNA polymerase according to supplier’s instructions (Promega). For sequencing with high G+C content, the reactions were performed at 75 °C and gel compressions minimized by lyophilizing the reaction mixtures before addition of 80% formamide gel loading buffer. Sequencing reactions were heated in boiling water and analysed on 4% or 7% polyacrylamide gels containing 8 M urea (Sambrook et al., 1989).

Southern blotting. DNA from agarose gels was transferred to Zeta-Probe nylon hybridization membrane (Bio-Rad Laboratories) using either capillary transfer or pressure blotting. DNA was fixed to the membrane by baking at 80 °C for 20 min in vacuo. Membranes were prehybridized in 5–10 ml 50% formamide, 5 × SSC, 5 × Denhardt’s solution, 5 mM sodium pyrophosphate, 250 μg denatured and sheared herring sperm DNA ml⁻¹, 0.5% SDS for 1–12 h at 42 °C. Hybridization was performed in a solution of 50% formamide, 5 × SSC, 1 × Denhardt’s solution, 5 mM sodium pyrophosphate, 100 μg denatured and sheared herring sperm DNA ml⁻¹, 0.5% SDS and ³²P-labelled RNA probe for 18–21 h at 42 °C. Blots were washed twice for 10 min each at 75 °C in 0.1 SSC, 0.1% SDS followed by a 20 min wash at 65 °C in 0.1 SSC, 0.1% SDS.

Construction of deletion subclones. The 1.9 kb insert of GLO7 was subcloned into the SalI site of Bluescript SK+ (Stratagene) and a clone, SKG7, was selected in which transcription from the Frankia Arl5 DNA was in the opposite direction to that of the vector β-galactosidase gene. SKG7 DNA was digested with XhoI/KpnI and nested deletions were created using exonuclease III with S1 nuclease.

DNA extraction and RT-PCR. These procedures were done as described (T. John, J. Rice & J. Johnson, unpublished) except strain Arl5 cells grown as described in bacterial cultures were used. Primers for PCR were 5’-CACGCGTCAATCGCTTTCTTC-3’ and 5’-GAAAACGTACCCCTGCTGG-3’, which anneal to positions 1000–1018 and 1320–1301 from the GTG start codon (see Fig. 5a).

RESULTS

Activity of Frankia DNA–luciferase fusions in E. coli

The search for an endogenous Frankia promoter was undertaken utilizing luciferase promoter-probe clones as described in Methods. Since a subset of Streptomyces promoters are known to be active in E. coli (Bibb & Cohen, 1982; Buttner & Brown, 1987), constructs were initially transformed into E. coli and 2000–3000 colonies were visually screened for bioluminescence (Legocki et al., 1986). Twelve of the clones exhibited bioluminescence above background levels. Cells from these GLO clones were assayed in a luminometer and given numbers that reflected their relative luminescence, and therefore promoter activity, in E. coli (Fig. 1).

Transcription activity of GLO clones in Frankia extracts

To assay the GLO clones for promoter activity in Frankia, in vitro transcription reactions were performed as described in Methods. Four clones, GLO2, 3, 4 and 7, exhibited transcription activities at least threefold greater than that of the ∆P template based on TCA precipitation assays (data not shown). To determine the extent of lux transcription from each of the GLO clones,
radionucleotide transcripts produced in the *Frankia* extract were hybridized to Southern blots of their respective template DNAs that had been digested with *SalI/BamH*I. This digestion produced a 4–4 kbp fragment derived mainly from pBR322 sequence, a 3–4 kbp *luxAB* fragment and a *Frankia* insert fragment unique to each clone. Under RNA-limiting conditions, the intensity of hybridization to the *luxAB* band is a measure of the level of transcription from any correctly oriented promoter in the *Frankia* insert. The relative intensity of the insert band should be proportional to the distance of the promoter element from the *luxAB* proximal *SalI* site. As an approximate measure of promoter activity, we determined the extent of hybridization to the *luxAB* band normalized to the pBR322 band. This accounts for transcription that may initiate within the pBR322 fragment and continue into the *luxAB* region. If transcription reads through from *luxAB* into pBR322 sequences, or if there is any transcription initiated in the opposite direction, the level of *luxAB* initiation within the *Frankia* insert will be underestimated.

The inserts in GLO3 and GLO7 exhibited significantly higher *lux* transcription in the *Frankia* extract than pFIT001 (Fig. 2). All GLO clones produced more *luxAB* transcript than the promoter-deleted clone, ΔP. In addition, GLO4 and GLO7 produced RNAs that hybridized to the insert bands, suggesting that a transcription start site was present within the insert.

**Identification of transcription start sites in GLO clone inserts**

To detect *lux* transcripts that initiated within the insert fragments of GLO2, 3, 4 and 7, an oligodeoxynucleotide complementary to a region of *luxAB*, the ΔP primer, was hybridized to *in vitro* transcripts produced in the *Frankia* extract and extended with MMLV reverse transcriptase. Extension products were analysed by electrophoresis in a denaturing polyacrylamide gel along with dideoxynucleotide sequencing reactions of the GLO7 clone using the same primer. Only the GLO7 clone transcripts yielded detectable reverse transcriptase bands. Two products were visible from GLO7 with lengths of 217 and 166 bp, designated as transcription start points (TSPs) 1 and 2, respectively (Fig. 3). Clone GLO3 did not produce a detectable primer extension product even though, in the previous hybridization analysis, it exhibited greater *luxAB* transcription than GLO7. The intensity of the GLO3 insert band in Fig. 2 suggests that the promoter activity may be located too far upstream for the primer to be completely extended by reverse transcriptase.

**Localization of promoter sequences in GLO7**

The *Frankia* insert of GLO7 was excised and subcloned into pBluescript SK +, producing SKG7. The putative *Frankia* promoter element was oriented in the opposite

---

**Fig. 2.** Direction of transcription from GLO clone promoter sequences. Plasmid DNAs (0.5 µg) were digested with *SalI/BamH*I, separated by agarose gel electrophoresis and transferred to nylon membranes. Transcripts (1–2 × 10⁶ d.p.m. each) were produced as described in Methods, purified by organic extractions, ethanol precipitated, then hybridized to blots of cognate DNA. An autoradiograph is shown. ∆P, *E. coli* DNA digested with *HindIII*. The extent of hybridization to *luxAB* sequence was determined by densitometer analysis. An autoradiographic exposure was made of the blot in which the absorbance readings were within the experimentally determined linear range of the film and densitometer. The numbers at the bottom are the ratios of the intensities of the *luxAB* bands to the cognate pBR322 band.
Frankia promoters

Fig. 3. Primer extension analysis of in vitro transcripts from GLO clones. Non-radiolabelled transcripts, prepared as described in Methods, were used as templates for primer extension reactions with ³²P-labelled ΔP primer as described in Methods. Products were analysed on a 7% polyacrylamide gel containing 8 M urea. Sequencing reactions performed on the GLO7 clone using the same primer are included as size markers.

Fig. 4. Primer extension analysis of in vitro transcripts from SKG7 deletion clones. Transcription reactions were performed using 16 µg Frankia extract protein and 1.0 µg DNA from each of the SKG7 deletion clones Δ1–Δ6 or pBluescript SK+ as the template. Primer extension reactions were performed on the transcripts using ³²P-labelled T3 primer as described in Methods. Products were analysed on a 7% polyacrylamide gel containing 8 M urea. Sequencing reactions performed on the SKG7 clone with the same primer are included as size markers.

direction to the β-galactosidase promoter to eliminate the possibility of interference in transcription and primer extension analyses. Incremental deletions of the SKG7 insert were produced from the 5′ side of the transcription start sites using exonuclease III plus S1 nuclease. The end points of six clones were determined by sizing of insert bands on agarose gels (SKG7Δ1 and Δ2) or DNA sequencing (SKG7Δ3–Δ6). The end points of deletion clones Δ1 to Δ6, relative to the location of TSP-1, were −950, −450, −324, −176, −55 and +39, respectively. Primer extensions using end-labelled T3 primer were performed on in vitro transcripts made from SKG7, deletion clones Δ1–Δ6 and pBluescript SK+ using the Frankia extract. Extension products were analysed by electrophoresis in a denaturing polyacrylamide gel along with oligodeoxynucleotide sequencing reactions of SKG7 from the same primer. The major primer extension products from the deletion clones were identical to the TSP-1 and TSP-2 previously identified, with TSP-2 being more abundant (Fig. 4). The TSP-1 product was produced from all clones except Δ6, which does not contain this transcription start site. The TSP-2 product was evident in all lanes except the vector control.

The presence of the TSP-1 band in Δ5 localized essential promoter element(s) to within 55 bases upstream of TSP-1. The Δ6 clone produced a primer extension product corresponding to TSP-2 even though it contained only 12 bp upstream of TSP-2 from the original GLO7 clone. These regions are referred to as GLO7p1 and GLO7p2, respectively. The band migrating more slowly than TSP-2 in the Δ6 lane appears to have initiated within a portion of pBluescript sequence that exhibits no significant homology to the GLO7p1 or GLO7p2 sequences. A corresponding primer extension band was not produced from the other GLO clones or from pBluescript alone.

Comparison of GLO7 with other bacterial promoter elements

The sequence of the luxAB proximal region of the GLO7 insert was determined from GLO7 and SKG7Δ3–Δ6. An overlapping clone, produced by BclI digestion of Frankia ArI5 DNA, was isolated using the SKG7Δ6 fragment as a probe. The combined sequence is deposited in GenBank under accession no. AY008259. The locations of TSP-1 and 2 as well as the end points of SKG7 clones
sequences at the 5’ end of the DNA sequence. (b) The sequence of GLO7p1 aligned with the TSP-1 and TSP-2 regions that deviate from the consensus are presented with gaps introduced to maximize identity. Capital letters represent nucleotides in the DNA sequence. Regions with the best fit to the TSP-1 and TSP-2 sequences are identified by shadows. (c) The sequences of GLO7p2 and GLO7 derived sequences in SKG7 are identical to the GLO7p1 sequence. TSP-1 is indicated by an arrow. The positions of TSP-1 and TSP-2 are identified by a triangle. The sequences of GLO7p2 and SKG7Δ6 are identical to the GLO7p1 sequence. TSP-1 is indicated by an arrow. The sequences of GLO7p2 and SKG7Δ6 are aligned with the consensus nif promoter sequence (Barrios et al., 1999). Capital letters represent nucleotides identical to the nif promoter. The junction of pBluescript SKI1 and GLO7 derived sequences in SKG7Δ6 is marked with an arrow.

Δ5 and Δ6 are identified in Fig. 5(a). The GLO7 sequence was searched for similarity to promoter elements identified in other bacteria. A sequence similar to the NtrC consensus binding site identified in E. coli and Salmonella (Buck et al., 1986) and to a similar element upstream of the Frankia CpII glnII gene (Rochefort & Benson, 1990) was found beginning 101 bp upstream of TSP-1 (Fig. 5b).

In addition, the immediate upstream regions of both TSP-1 and TSP-2 contained sequences with identity to E. coli canonical promoter elements (Fig. 5a). These similarities are apparently sufficient for recognition in E. coli as evidenced by primer extension of RNA transcribed in vitro by purified E. coli RNA polymerase from the GLO7 DNA as template. Two primer extension products corresponding exactly to TSP-1 and TSP-2 were present, but at much lower levels than those produced from RNA made by Frankia extracts (data not shown). A region with 5 of 7 identity to a Streptomyces nif − 35 element (Taguchi et al., 1989) is also present in an appropriate region upstream of TSP-1 (Fig. 5a).

The TSP-1 region also has two sequence blocks, centred at −8 and −29 bp upstream, that have substantial identities at or near the 5’ end of the Frankia CpII glnII gene (Fig. 5b). An element similar to the nif/ntr promoter consensus characterized in Rhizobium and Klebsiella (Ausubel, 1984; Barrios et al., 1999) is present between 2 and 17 bp upstream from TSP-2 (Fig. 5c).

In vivo activity of the TSP-1 and 2 sequences

The GLO7 fragment terminated 136 bp downstream from TSP-1. A clone overlapping GLO7 by 853 bp was isolated from a BclI digest of Frankia ArI5 DNA. Sequence analysis of this clone revealed the presence of an ORF encoding 498 aa that initiated with a GTG codon 64 bp downstream from TSP-1 (Fig. 5a). A Shine–Dalgarno sequence, GGAGG, begins 12 bp upstream of the putative start codon, 3 bp beyond the usual range of spacing (Barrick et al., 1994). The ORF has a high G+C content in the third codon position (76-4%) and is terminated by tandem UGA codons (data not shown). All of these features suggest that this ORF encodes a protein. However, a BLASTP comparison of the deduced amino acid sequence of the putative protein with the GenBank database does not identify significant identities with any known protein. The best matches, E = 2 x 10−3, are with other hypothetical proteins rich in proline and arginine.

In vivo transcription of the ORF was verified by RT-PCR using total RNA extracted from Frankia ArI5 cells. The RNA preparation was used as template with random decanucleotide primers to produce a cDNA copy. This DNA was amplified by PCR using primers nested within the ORF. A fragment of the expected size, 321 bp, was produced in the reaction (data not shown). No such product was evident in a control reaction in which the reverse transcriptase step was omitted. The product is therefore dependent on RNA present in the cell extract. The DNA sequence of the 321 bp RT-PCR fragment corresponds perfectly to the region of the ORF bracketed by the primers identified in Methods. This result supports the conclusion that the promoter region(s) identified by in vitro analyses also function in vivo.

DISCUSSION

In an attempt to elucidate sequences involved in transcriptional regulation by the symbiotic diazotroph Frankia, a promoter probe vector was used to identify Frankia DNA fragments that stimulated transcription in E. coli (Fig. 1). These clones were then used as templates for a Frankia cell-free transcription extract to determine whether any of the Frankia fragments could also direct transcription in the homologous system (Fig. 2). Primer extension analysis of RNA produced in vitro identified...
two transcription start sites, TSP-1 and TSP-2, that are 52 bp apart (Fig. 3). In these extracts, the upstream sequences necessary for transcription at TSP-1 have been localized by deletion analysis to within 55 bases of the start site (Fig. 4). This region, referred to as GLO7p1, contains elements similar to E. coli canonical —35 and —10 sequences (Fig. 5a), and to the Streptomyces ssi gene —35 element (Obata et al., 1989; Taguchi et al., 1989) (Fig. 5a). The glnII gene of Frankia strain CpIII contains a striking homology to GLO7p1 in a region tentatively identified as encoding the N terminus of the protein (Rochefort & Benson, 1990). An alignment of the Frankia sequences reveals two regions of complete identity: a —35 element, GAGTT, and a —10 region, GATCG (Fig. 5b). These are followed by CGTTT and either TTTT (GLO7p1) or CACC (glnII). In the GLO7 clone, both TSPs initiate with the G immediately preceding the polypyrimidine sequence. Since neither the transcriptional nor translational start sites were experimentally determined for glnII, it is possible that this region is used as a promoter and translation initiates downstream. Further analysis of glnII transcription is needed to resolve this uncertainty. The sequence necessary for transcription at TSP-2 in these extracts has been shown by deletion analysis to be limited to 12 bp upstream of the start site (shown by deletion analysis to be limited to 12 bp for transcription at TSP-2 in these extracts has been needed to resolve this uncertainty. The sequence necess-

like protein, that the smaller mRNA may be transcribed by an NtrA-nif gene —35 element (Obata et al., 1989; Taguchi et al., 1989) (Fig. 5a). The glnII gene of Frankia strain CpIII contains a striking homology to GLO7p1 in a region tentatively identified as encoding the N terminus of the protein (Rochefort & Benson, 1990). An alignment of the Frankia sequences reveals two regions of complete identity: a —35 element, GAGTT, and a —10 region, GATCG (Fig. 5b). These are followed by CGTTT and either TTTT (GLO7p1) or CACC (glnII). In the GLO7 clone, both TSPs initiate with the G immediately preceding the polypyrimidine sequence. Since neither the transcriptional nor translational start sites were experimentally determined for glnII, it is possible that this region is used as a promoter and translation initiates downstream. Further analysis of glnII transcription is needed to resolve this uncertainty. The sequence necessary for transcription at TSP-2 in these extracts has been shown by deletion analysis to be limited to 12 bp upstream of the start site (Δ6 in Fig. 5a). This region, referred to as GLO7p2, contains a portion of the sequence homologous to the nif/ntr promoter (Fig. 5c). The TSP-2 upstream sequence matches the nif/ntr promoter consensus at 11 of 13 nucleotides, implying that the smaller mRNA may be transcribed by an NtrA-like protein, σ70-RNA polymerase complex. The Δ6 sequence is identical to GLO7p2 in the 3′-most 11 bases, which constitute the last three bases of the —35 region and all of the —10 region (Fig. 4). This may be sufficient to direct transcription although the invariant G residues are missing (Fig. 5c) (Barrios et al., 2000). The presence of an NtrC binding site approximately 145 bp upstream also suggests that transcription at GLO7p2 may be nitrogen regulated. It may seem contradictory to assert that GLO7p2 is an ntr promoter when in vitro transcription did not depend upon upstream sequences beyond —12, including the proposed NtrC binding site. However, the polymerase extracts used in these experiments were produced from cells grown in the presence of NH4Cl that, by analogy to E. coli, would contain only small amounts of active NtrC. The presence or absence of an NtrC binding site would therefore not have been expected to result in changes in transcriptional activity. In addition, it has been shown that σ70-RNA polymerase transcription can be activated by other proteins that affect DNA conformation (Buck et al., 2000).

Because the promoters described here are located at the extreme 3′ end of the GLO7 insert, the identity of any corresponding gene was not evident. An overlapping fragment of Frankia ArI5 DNA was therefore cloned to determine whether the putative promoter region was associated with any functional gene(s). The proximal ATG downstream from GLO7p2 is in-frame with a stop codon. DNA sequence analysis of the overlapping fragment identified a 498 aa ORF with a potential GTG start codon 64 bp downstream from TSP-1. The ORF has a Shine–Dalgarno sequence (Gold, 1988) and bias in third codon position G + C content of 76.4%. Further, the ORF is terminated by tandem nonsense codons. These are all indicators that the ORF is a functional gene.

To determine whether this region of the Frankia DNA is transcribed in vivo, an RNA preparation from actively growing ArI5 cells was used as the template for RT-PCR. The PCR primers were nested within the ORF and produced a 321 bp fragment, as expected from the corresponding DNA sequence (data not shown). A control reaction in which the reverse transcription was not done showed no product, indicating that the PCR depends on an RNA template. The DNA sequence of the amplified fragment was identical to the ORF, confirming that the PCR product represents transcription from the ORF. The presence of the in vivo transcript provides strong support for the conclusion that the transcription analyses done in vitro reflect the in vivo activity of the Frankia RNA polymerase.

Most Streptomyces promoters are not recognized in E. coli, perhaps due to differences in G + C content (Bibb & Cohen, 1982; Buttner & Brown, 1987). In this study, Frankia promoters that are operative in E. coli were selected by a promoter screening in that organism. The promoters GLO7p1 and GLO7p2 are homologous to E. coli canonical promoter elements and the transcription start sites were shown to be identical using both Frankia and E. coli RNA polymerase in vitro. Whether or not recognition in E. coli is a general feature of Frankia promoters awaits further investigation.

The presence of sequences resembling NtrC binding sites (Fig. 5a) and σ70-RNA polymerase recognition sequences (Fig. 5b,c) in Frankia suggests that the actinomycete may regulate ntr genes in a manner similar to other diazotrophs and that it may contain proteins that are functionally and structurally similar to NtrA and NtrC. However, this conclusion must be tempered by the lack of similarity of the downstream ORF to any known ntr gene.

ACKNOWLEDGEMENTS

J. V. B. was supported by a Patricia Roberts Harris Fellowship. The work was supported by the Agricultural Experiment Station of the University of Wyoming and NSF awards RII-8610680 and NSFLOC4510 to J. D. J.

We gratefully acknowledge the contributions of Ms Shawna Clark, Ms Holly Magna and Ms Katja Manninen to the success of this work.

REFERENCES


Reitzer, L. & Magasanik, B. (1986). Transcription of glnA in E. coli is stimulated by activator bound to sites far from the promoter. Cell 45, 785–792.


Received 19 July 2000; revised 3 October 2000; accepted 16 October 2000.