Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 376 (2006) 144–151
GENE
SECTION
FUNCTIONAL GENOMICS
www.elsevier.com/locate/gene

Identification and characterization of the gene encoding the *Acidobacterium capsulatum* major sigma factor

Zakee L. Sabree [a,b], Veit Bergendahl [c,1], Mark R. Liles [a,2], Richard R. Burgess [c], Robert M. Goodman [a,3], Jo Handelsman [a,*]

[a] *Department of Plant Pathology, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA*
[b] *Microbiology Doctoral Training Program, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA*
[c] *McArdle Laboratory for Cancer Research, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA*

Received 11 December 2005; received in revised form 14 February 2006; accepted 15 February 2006
Available online 5 April 2006
Received by R. Britton

Abstract

Acidobacterium capsulatum is an acid-tolerant, encapsulated, Gram-negative member of the ubiquitous, but poorly understood Acidobacteria phylum. Little is known about the genetics and regulatory mechanisms of *A. capsulatum*. To begin to address this gap, we identified the gene encoding the *A. capsulatum* major sigma factor, *rpoD*, which encodes a 597-amino acid protein with a predicted sequence highly similar to the major sigma factors of *Solibacter usitatus* Ellin6076 and *Geobacter sulfurreducens* PCA.

Purified hexahistidine-tagged RpoD migrates at ~70 kDa under SDS-PAGE conditions, which is consistent with the predicted MW of 69.2 kDa, and the gene product is immunoreactive with monoclonal antibodies specific for either bacterial RpoD proteins or the N-terminal histidine tag. *A. capsulatum* RpoD restored normal growth to *E. coli* strain CAG20153 under conditions that prevent expression of the endogenous *rpoD*. These results indicate we have cloned the gene encoding the *A. capsulatum* major sigma factor and the gene product is active in *E. coli*.

Keywords: Acidobacteria; *rpoD*; Inverse PCR; Complementation; Heterologous expression

1. Introduction

Acidobacterium capsulatum ATCC 51196 is a high G+C, acidophilic, Gram-negative rod originally isolated from acid mine drainage (Kishimoto et al., 1991). *A. capsulatum* is a member of the Acidobacteria phylum subgroup 1, which includes 16S rRNA gene sequences and only a few cultured members (Joseph et al., 2003; Davis et al., 2005). Multiple culture-independent microbial community analyses indicate that the Acidobacteria are ubiquitous and abundant across a wide variety of soil types, suggesting that they play important roles in soil microbial communities (Barns et al., 1999). Abundance in the environment is also indicated in metagenomic libraries that contain DNA cloned directly from soil (Liles et al., 2003; Quaiser et al., 2003). Although the Acidobacteria have been described phylogenetically, the mechanisms by which they regulate gene expression remain uncharacterized. In bacteria, most housekeeping genes are regulated by the major sigma factor, RpoD, which coordinates initiation of RNA polymerase (RNAP)-dependent transcription through promoter recognition and DNA melting (Burgess et al., 1969; Gross et al., 1998). Although all bacterial major sigma factors share this regulatory function, the breadth of promoter specificity differs substantially between major sigma factors from phylogenetically

Abbreviations: Km, kanamycin; Cm, chloramphenicol; GYE, glucose-yeast extract; LB, Luria–Bertani medium; LBA, Luria–Bertani agar; IPTG, isopropyl thiogalactoside; dNTP, deoxynucleoside triphosphates.

* Corresponding author. Department of Plant Pathology, Russell Laboratories, 1630 Linden Drive, Rm. 583, Madison, WI 53706, USA. Tel.: +1 608 263 8783; fax: +1 608 265 5289.
E-mail address: joh@plantpath.wisc.edu (J. Handelsman).
[1] Present address: Genetics/Biotechnology Center 4422, University of Wisconsin-Madison, Madison, WI 53706, USA.
[2] Present address: Biological Sciences, Auburn University, Auburn, Alabama 36849, USA.
[3] Present address: Rutgers' Cook College/NJAES Rutgers, The State University of New Jersey, New Brunswick, NJ 08901, USA.

0378-1119/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2006.02.033

distinct bacteria (Browning and Busby, 2004; Gruber and Bryant, 1997). In this paper, we report that we have cloned the major sigma factor from *A. capsulatum*, purified its gene product, and demonstrated that *A. capsulatum rpoD* restores normal growth to an *rpoD* mutant of *E. coli* under growth-restrictive conditions.

2. Materials and methods

2.1. Bacterial strains and plasmids

Bacterial strains and plasmid constructs used in this study are shown in Table 1. *A. capsulatum* ATCC 51196 was grown at 28 °C and maintained in glucose-yeast extract (GYE) media (Kishimoto et al., 1991) or maintained on GYE media, pH 3.5, +8 g/L Gelrite gellan gum. *E. coli* strains used for cloning and protein expression were grown at 37 °C and maintained in Luria–Bertani (LB) broth or agar (LBA) supplemented with 100 μg/mL of ampicillin, 50 μg/mL of kanamycin (Km), 25 μg/mL of chloramphenicol (Cm), 200 μM indole acrylic acid, 1 mM IPTG, and 33 mM arabinose as required.

2.2. Genomic DNA and plasmid isolation

Two milligrams per milliliter of genomic DNA was isolated from 100 mg (wet weight) *A. capsulatum* cell pellet using the EasyDNA kit (Invitrogen). Plasmids were isolated using QIAprep Spin Miniprep kits (QIAGEN). DNA isolation kits were used according to the manufacturers' protocols.

2.3. Degenerate PCR

The amino acid sequences of the major sigma factors from over 60 phylogenetically diverse prokaryotic species were aligned using MegAlign (DNAStar) to identify highly conserved regions. Primers (dsigmag F′: 5′-AGCTTTGGTGGATYMGNCARGC-3′ and reverse sigma2_R′: 5′-TCGATCTGACGAA-TACGTTCGCGGGTAAC-3′) were designed using the conserved amino acid sequence from Region 2.4 (WWIRQA) and Region 4.2 (VTRERIRQIE), respectively, to amplify a nested segment of the *A. capsulatum rpoD* by PCR (1× 95 °C for 2 min; 25× 95 °C for 30 s, 60 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min; 1× 72 °C for 5 min).The PCR reaction contained: 0.5 U Taq polymerase, 2× reaction buffer, 0.2 mM concentration of all four deoxynucleoside triphosphates (dNTPs), and 0.75 μM forward and reverse primers in a 25-μL reaction volume. Standard gel electrophoresis was used to confirm amplification of a 476-bp PCR product that was subsequently cloned into pGEM-T to give pZLS1, using a pGEM-T cloning kit (Promega), and transformed into *E. coli* DH5α by electroporation (Ausubel et al., 1994). Nucleotide sequencing was performed by a capillary automated DNA cycle sequencing method using fluorescently labeled dideoxy terminators and M13 primers (Integrated DNA Technologies) by the University of Wisconsin Biotechnology Center (http://biotech.wisc.edu/).

2.4. Cloning A. capsulatum rpoD

Degenerate inverse PCR (Fig. 1) was used to amplify DNA flanking the 476-bp PCR product. DNA template for inverse PCR was prepared by treatment of 10 μg of *A. capsulatum* genomic DNA with 20 U of *Apa* I at 37 °C for 3 h, followed by an ethanol precipitation. *Apa* I restriction endonuclease was chosen because its recognition sequence was not found within the coding sequence of the major sigma factors from 50 phylogenetically diverse bacteria or the 476-bp fragment by MapDraw (DNAStar) analysis (data not shown). One microgram of *Apa* I-treated DNA was recircularized with T4 DNA ligase (4 U) at 25 °C for 72 h followed by ethanol precipitation. Nested primers used in inverse PCR were derived from the 476-bp product previously amplified from *A. capsulatum* genomic DNA. Each 25-μL PCR mixture contained 100 ng of self-ligated *A. capsulatum* genomic DNA, 0.75 μM forward and reverse primers (CON_SIG_F3 F′: ACACGCTC-GAAGAAGTAGGCCAGA and CON_SIG_R3 R′: TGATCGTCTCAATCATGTGCACCGGAATA) in addition to the previously described quantities of PCR buffer, dNTPs and *Taq* polymerase. DNA amplification conditions were as follows: (1× 95 °C for 1 min; 25× 95 °C for 1 min, 65 °C for 30 s, 59 °C for 45 s,

Table 1
Bacterial strains and plasmids

Bacteria strain or plasmid	Genotype	Reference
Acidobacterium capsulatum ATCC 51196		Kishimoto et al.
E. coli DH5α	F^- φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *deoR, recA1 endA1 hsdR17* (r_k^- m_k^+ *phoA supE*44 λ^- *thi*-1 *gyrA*96 *relA*1	Invitrogen
E. coli BL21 Star	F^-*ompT hsd*S_B ($r_B^-m_B^-$) *gal dcm rne*131 (DE3)	Invitrogen
ACRPOD1	*E. coli* BL21 Star pZLS4	This paper
BL21-pET28b	*E. coli* BL21 Star pET28b	This paper
pGEM-T	pGEM5Z-F(+) Amp^r *lacZ*	Promega
pZLS1	pGEM-T+476 bp *A. capsulatum* DNA-derived PCR amplified product	This paper
pZLS2	pGEM-T+2.4 kb *A. capsulatum* DNA-derived PCR amplified product	This paper
pZLS3	pGEM-T+1794 bp *A. capsulatum rpoD*	This paper
pET28b	P_{lac}-T7, Kan^r	Novagen
pZLS4	pET28b+*A. capsulatum rpoD*	This paper
pCF430	P_{ara} Tet^r	Newman et al.
pCF430AcRpoD	pCF430+*A. capsulatum rpoD*	This paper
CAG20153	*E. coli* C600K-Ω Cm^R P_{trp}-*rpoD zgh*: :Tn10	Lonetto et al.

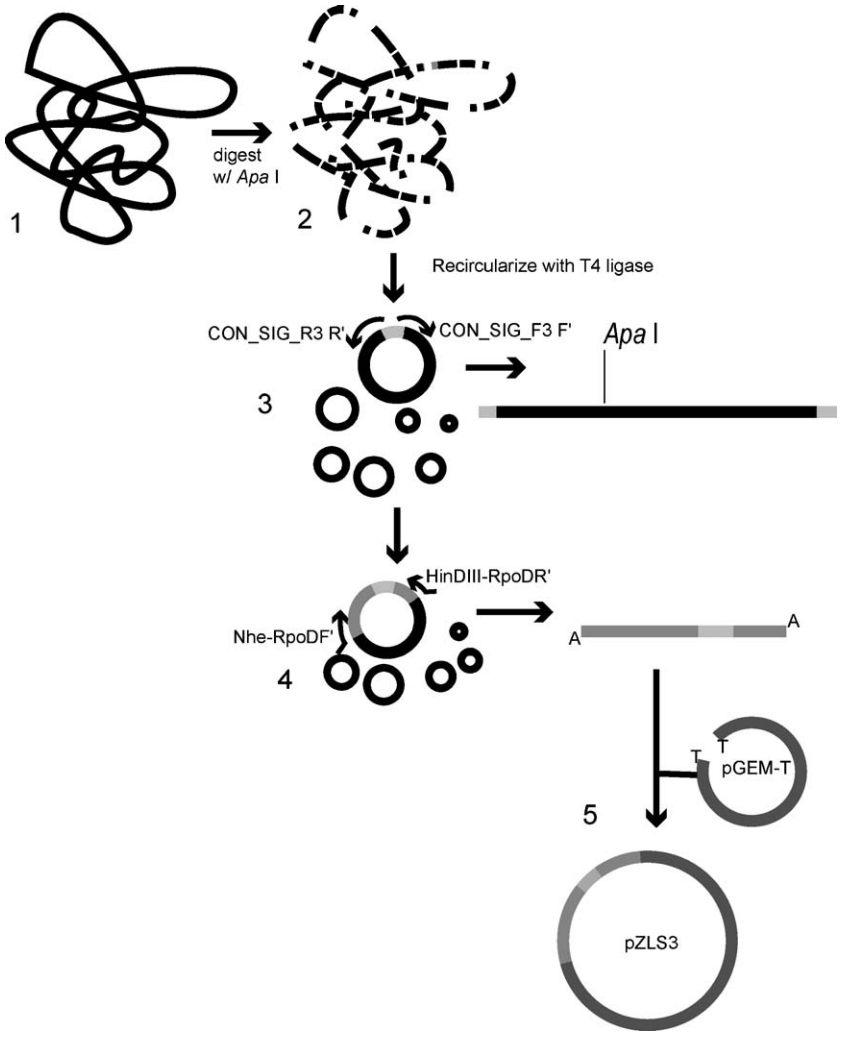

Fig. 1. Inverse PCR method. Genomic DNA was digested with *Apa* I (1) and self-ligated with T4 DNA ligase (2). The first round of PCR used primers based on a nested (light grey) region of *A. capsulatum rpoD*, which were used to amplify flanking region from the recircularized genomic DNA (3). Sequence of the PCR product revealed an *Apa* I site, which was used as a point of reference to determine that the genomic DNA fragment containing *rpoD* had been completely sequenced. A second round of PCR involved *Hin*D III- and *Nhe* I-linked primers designed to amplify *rpoD* (dark grey) using either self-ligated or genomic DNA template (4). PCR product was TA-cloned into pGEM-T to produce pZLS3 (5).

72 °C for 4 min; 1× 72 °C for 5 min). A single 2.4-kb PCR product was cloned into a pGEM-T plasmid, pZLS2, and sequenced. A final round of PCR was performed to amplify a 1794-bp ORF using 100 ng of self-ligated or untreated *A. capsulatum* genomic DNA, *Nhe* I-linked forward (Nhe-RpoDF′: GCTAGCCTTTGGGGGAAGGTG) and *Hin*D III-linked reverse (HinDIII-RpoDR′: AAGCTTTTGCTTACATTTCGCGGGC) primers in a 25-μL reaction under the following amplification conditions (1× 95 °C for 1 min; 25× 95 °C for 30 s, 55 °C for 1 min, 72 °C for 2 min; 1× 72 °C for 5 min). PCR product was cloned into pGEM-T, resulting in pZLS3, and sequenced using M13 primers.

2.5. Sequence alignment and analyses

Nucleotide sequence editing and PCR product alignments were carried out using Chromas version 2.3 (www.technelysium.com.au) and SeqMan (DNAStar). Nucleotide homology and deduced protein sequence homology and similarity searches were performed using BLASTN and BLASTX (http://www.ncbi.nlm.nih.gov/BLAST/) search engines, respectively (Altschul et al., 1997). Web-based CLUSTALW (http://www.ebi.ac.uk/clustalw/) program (Thompson et al., 1994) was used to conduct multiple sequence alignments with the primary sigma factors from the following bacteria as references (GeneBank accession numbers): *E. coli* RpoD (P00579), *Caulobacter crescentus* RpoD (AAK25009), *Bacillus* sp. C-125 SigA (AB012852), and *Solibacter usitatis* sigma-70 Ellin6076 (ZP_00527478). The unfinished *A. capsulatum* ATCC 51196 genome sequence is accessible for BLAST search at the website http://tigrblast.tigr.org/ufmg/. Conserved protein motif searches were performed using the NCBI Conserved Domain Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Marchler-Bauer and Bryant, 2004).

2.6. Overexpression and purification of A. capsulatum RpoD

A *Nhe* I–*Hin*D III-linked fragment excised from pZLS3 by double restriction digest with *Nhe* I and *Hin*D III was

unidirectionally ligated into identically treated pET28b (Novagen) downstream of the T7 promoter and hexahistidine tag to give pZLS4. Ten nanograms of pZLS4 DNA was transformed into 50 μL of competent *E. coli* BL21-Star cells (Invitrogen) and transformants were selected on LB agar+Km (40 μg/mL) incubated overnight at 37 °C. Transformants were rapidly screened for the presence of the insert by colony PCR using the Nhe-RpoDF′ and HinDIII-RpoDR′ primers in 25-μL reactions. Purified *E. coli* BL21(DE3) genomic DNA, pET28b plasmid DNA and *E. coli* BL21(DE3) pET28b were included as controls and were consistently negative in these screens (data not shown). A positive clone, ACRPOD1, was selected for plasmid preparation using the QIAprep kit to reconfirm the presence of the insert by double digest with *Hin*D III and *Nde* I.

Overexpression of RpoD was achieved by addition of 1 mM IPTG (final concentration) to 1 L mid-log phase ACRPOD1 LB+Km liquid culture (OD_{600}=0.6) followed by >5 h incubation at 37 °C. SDS-PAGE analysis of previous small-scale overexpression of *A. capsulatum rpoD* gene product localized the protein in the insoluble phase. Higher concentrations of IPTG did not result in higher amounts of protein, nor did lower IPTG concentrations, alterations of media pH or growth temperatures improve protein yield or solubility (data not shown). We prepared 1-ml aliquots of inclusion bodies from

```
                                          -35                   -10      rbs
     cgggaaaccaacgcgtcgtaccaGTGAAGacccctcgcatgatactcTATAAGtctttgggggaag

   1  V  R  P  S  S  R  D  F  V  R  N  G  T  C  L  S  W  P  R  S  H  F  V  R  F  N  A  R  S  D
   1 GTGCGCCCATCGTCCCGCGACTTTGTTCGCAATGGGACATGCCTGTCATGGCCTCGATCGCACTTCGTCCGATTCAACGCCAGATCCGAC

  31  H  I  R  G  R  R  I  R  L  A  L  D  D  K  Y  E  E  D  L  Q  K  L  I  D  A  G  K  E  K  G
  91 CATATTCGCGGGAGACGAATCCGCTTGGCATTAGACGATAAATACGAAGAGGATCTGCAGAAGCTCATCGACGCCGGTAAGGAAAAGGGC

  61  Y  L  T  Y  N  E  V  N  D  L  I  P  G  D  M  T  S  P  E  D  L  D  D  L  M  T  T  I  G  T
 181 TATTTGACCTACAACGAGGTCAACGACCTGATCCCTGGCGACATGACTTCTCCAGAAGACCTGGACGACCTGATGACGACCATCGGGACC

  91  Q  G  I  D  I  L  E  G  S  P  K  L  G  N  D  G  F  D  D  E  D  S  D  D  V  E  L  D  L  T
 271 CAGGGAATCGACATTCTTGAAGGCAGTCCCAAGCTCGGCAACGATGGCTTCGACGACGAAGACAGCGACGACGTCGAGCTTGATCTTACC

 121  P  G  T  L  E  K  T  N  D  P  V  R  M  Y  L  R  E  M  G  T  V  P  L  L  T  R  E  G  E  V
 361 CCAGGAACGCTTGAAAAAACCAACGACCCCGTGCGCATGTATCTGCGCGAAATGGGCACCGTTCCGCTGCTCACACGCGAAGGCGAAGTA

 151  E  I  A  K  R  I  E  R  G  Q  L  R  V  M  K  A  V  S  R  S  P  I  V  I  R  E  I  I  A  L
 451 GAAATCGCCAAGCGCATTGAACGCGGCCAGTTGCGCGTGATGAAGGCTGTCTCGCGCTCGCCCATCGTGATTCGCGAGATCATCGCGCTC

 181  G  E  D  L  R  R  G  V  R  S  V  K  E  V  V  I  F  D  E  E  E  I  T  E  E  V  L  A  A  R
 541 GGTGAAGATCTGCGCCGCGGTGTGCGCAGCGTCAAGGAAGTGGTGATCTTCGACGAAGAAGAGATCACTGAAGAAGTACTGGCCGCCCGC

 211  V  R  Q  T  A  N  K  I  D  G  I  V  K  H  Q  K  K  I  Q  A  L  E  E  K  M  Q  A  A  K  S
 631 GTGCGCCAGACGGCCAACAAGATCGACGGCATCGTCAAGCACCAGAAGAAGATTCAGGCGCTCGAAGAGAAGATGCAGGCCGCCAAGAGC

 241  K  V  K  E  Q  R  K  L  R  W  L  I  A  R  E  M  V  F  I  T  R  I  V  R  D  L  K  Y  T  N
 721 AAGGTGAAGGAGCAGCGCAAGCTGCGCTGGCTCATCGCGCGCGAAATGGTCTTCATCACGCGCATTGTGCGCGACCTGAAGTACACCAAT

 271  Q  E  R  K  R  L  L  D  K  V  N  K  T  V  D  S  M  R  T  L  E  R  Q  I  R  S  L  D  Q  K
 811 CAGGAGCGCAAGCGCCTGCTCGACAAGGTCAACAAGACCGTCGATTCCATGCGCACGCTTGAGCGCCAGATTCGCTCTCTCGACCAGAAG

 301  Y  D  Q  S  K  S  E  E  L  K  K  E  Y  R  R  Q  Q  K  N  C  K  T  D  L  E  K  L  E  E  E
 901 TATGACCAGTCGAAGAGCGAAGAGCTCAAGAAGGAATACCGCCGCCAGCAGAAGAACTGCAAGACCGACCTCGAAAAGCTCGAAGAGGAA

 331  A  G  L  S  I  A  D  L  K  R  T  Q  R  E  I  I  Q  G  D  M  D  A  E  Q  A  K  R  E  L  I
 991 GCCGGACTCTCCATTGCCGACCTGAAGCGCACGCAGCGCGAGATCATTCAGGGCGACATGGATGCCGAGCAGGCCAAGCGCGAGCTGATT

 361  E  A  N  L  R  L  V  V  S  I  A  K  K  Y  T  N  R  G  L  Q  F  L  D  L  I  Q  E  G  N  I
1081 GAAGCCAACCTGCGCCTGGTGGTCTCGATTGCCAAGAAGTACACCAATCGCGGCCTGCAGTTCCTCGATCTGATTCAGGAAGGCAACATC

 391  G  L  M  K  A  V  D  K  F  E  Y  R  R  G  Y  K  F  S  T  Y  A  T  W  W  I  R  Q  A  I  T
1171 GGCCTGATGAAGGCCGTGGACAAGTTCGAGTACCGCCGCGGTTACAAGTTCTCCACCTACGCCACCTGGTGGATTCGCCAGGCCATCACG

 421  R  A  I  A  D  Q  A  R  T  I  R  I  P  V  H  M  I  E  T  I  N  K  L  I  R  T  S  R  Q  L
1261 CGCGCCATCGCCGATCAGGCCCGCACCATCCGTATTCCGGTGCACATGATTGAGACGATCAACAAGCTCATCCGCACCTCGCGTCAACTG

 451  V  Q  E  L  G  R  E  P  T  S  E  E  I  A  R  R  M  D  I  P  V  A  K  V  R  K  V  L  K  I
1351 GTGCAGGAGCTGGGCCGCGAACCCACCTCAGAGGAAATTGCCCGCCGCATGGATATTCCGGTGGCCAAGGTCCGCAAGGTGCTCAAGATT

 481  A  Q  E  P  I  S  L  E  T  P  I  G  E  E  E  D  S  H  L  G  D  F  I  E  D  S  Q  A  V  S
1441 GCGCAGGAGCCCATCTCCCTCGAAACGCCGATCGGCGAAGAGGAAGATTCGCACCTGGGCGACTTCATTGAGGACTCGCAGGCGGTCTCG

 511  P  S  E  A  V  I  S  V  N  L  K  E  Y  T  S  Q  V  L  R  T  L  T  P  R  E  E  R  V  I  K
1531 CCTTCAGAGGCGGTCATCAGCGTGAACCTCAAGGAGTACACCTCGCAAGTGCTGCGCACCCTGACACCGCGCGAAGAGCGCGTCATCAAG

 541  M  R  F  G  L  E  D  G  S  E  H  T  L  E  E  V  G  Q  S  F  Q  V  T  R  E  R  I  R  Q  I
1621 ATGCGCTTCGGCCTCGAAGACGGCTCAGAGCACACGCTCGAAGAAGTAGGCCAGAGCTTCCAGGTCACCCGCGAGCGCATTCGCCAGATC

 571  E  A  K  A  L  R  K  L  R  H  P  S  R  S  R  K  L  R  A  F  V  D  N  A  R  E  M  -
1711 GAAGCCAAGGCGCTGCGCAAGCTGCGGCACCCCTCGCGCTCGCGCAAGCTGCGGGCCTTTGTGGACAACGCCCGCGAAATGTAA

     gcaaaacagctacgaatgcaaAAAGGCCGCTCCCATCCGGAGCGGCCTTTgaggttgatgacaaacctat
```

Fig. 2. DNA sequence of *A. capsulatum rpoD*. Putative −35 and −10 regions are uppercased and indicated. A putative ribosome binding site is underlined and indicated. The coding region of *rpoD* is indicated by uppercase letters and the start codon is italicized. Deduced amino acids of the *rpoD* coding region are indicated by single letter abbreviations above their corresponding codons. The inverted repeat region of the Rho-independent terminator is indicated by underlined uppercase letters.

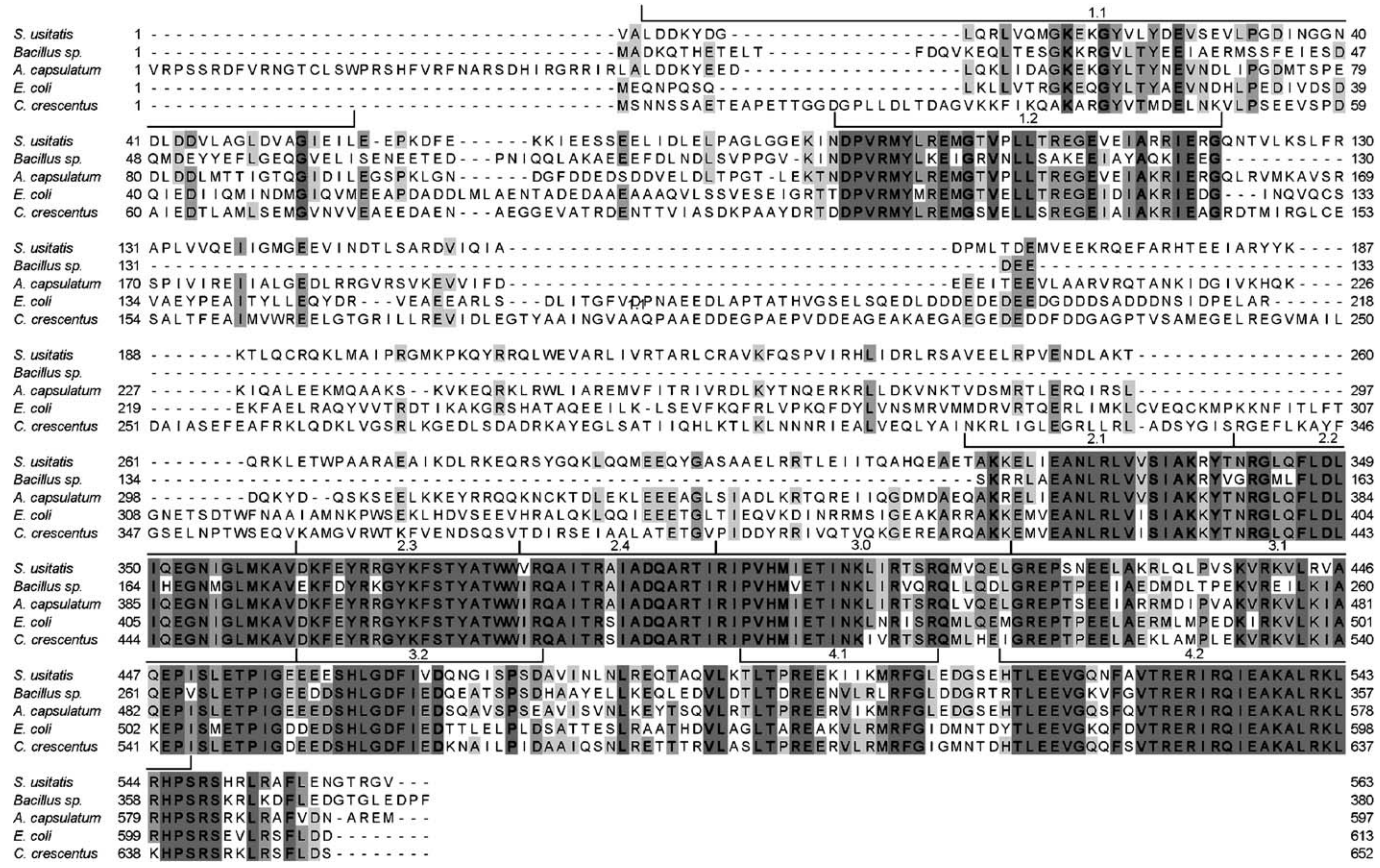

Fig. 3. Comparison of deduced amino acid sequences of bacterial major sigma factors. Sequence alignment was performed by CLUSTALW program (Thompson et al., 1994). Gaps, indicated by dashes, were introduced to optimize alignment. Conserved regions of the major sigma factors are indicated by the numbers above the sequences. Sequences used in these alignments are *E. coli* — *Escherichia coli* RpoD, *C. crescentus* — *Caulobacter crescentus* RpoD, *Bacillus* sp. — *Bacillus* sp. C-125 SigA, and *S. usitatus* — *Solibacter usitatis* sigma-70 Ellin6076, *A. capsulatum* — *Acidobacterium capsulatum*. Percent identity is highlighted as follows: dark grey — 100%, medium grey — 80% and light grey — 60%.

2.4 g of (wet weight) harvested cells using methods described by Bergendahl and Burgess (2003). One milliliter of inclusion bodies was solubilized in 10 mL NTGD buffer+6 M Gu-HCl and 5 mM imidazole and incubated for 30 min at room temperature. Following centrifugation to remove insoluble material, solubilized inclusion bodies were slowly dripped into 190 mL NTGD buffer+5 mM imidazole, stirred constantly at room temperature. Proteins were allowed to refold for 1 h at room temperature and centrifuged at 15,000 rpm for 10 min to remove precipitate.

Ni-NTA metal affinity chromatography was used to concentrate and purify refolded hexahistidine-tagged *A. capsulatum* RpoD according to methods accompanying the kit (QIAGEN). Fractions containing high quantities of protein, as determined by Bradford colorimetric assay, were concentrated using CentriCon (Millipore) concentrators and then diluted with 50% glycerol to a final concentration of 0.4 mg/mL and stored at −20 °C.

2.7. SDS-PAGE analyses and Western blot

Samples were taken throughout protein purification to track progress and were visualized by Laemmli electrophoresis methods (Ausubel et al., 1994) through pre-cast 4–12% SDS-PAGE gels (Invitrogen) using 1× MES buffer at 200 constant volts for 35 min. Gels were stained using Simply Stain (Invitrogen) Coomassie blue dye. For Western blots, proteins were transferred from gels to PVDF membranes using an XCell II Blot module with 1× transfer buffer (25 mM Bicine, 25 mM Bis–Tris, 1 mM EDTA)+10% methanol at 30 V for 1.25 h. Immunodetection was performed using the WesternBreeze (Invitrogen) Chromogenic kit with 2G10 (NeoClone) and anti-His (Novagen) primary antibodies diluted 1:2000 and 1:5000, respectively, in diluent supplied by the manufacturer. All protein gels and blotted membranes were scanned using a Hewlett-Packard ScanJet 3970.

2.8. Complementation of E. coli CAG20153

To determine whether the *A. capsulatum rpoD* gene product was functional in vivo in *E. coli*, we cloned *A. capsulatum rpoD* into pCF430 (Newman and Fuqua, 1999) at the *Nhe* I and *Hin*D III sites within the polylinker, resulting in pCF430AcRpoD. In *E. coli* CAG20153, the chromosomal *rpoD* promoter has been replaced with a *trp* promoter, placing it under the control of the *trp* repressor complex (Lonetto et al., 1998). CAG20153 *rpoD* is not expressed in the presence of tryptophan because the tryptophan-*trp* repressor complex blocks transcription, preventing cell growth. Addition of the tryptophan analog, indole

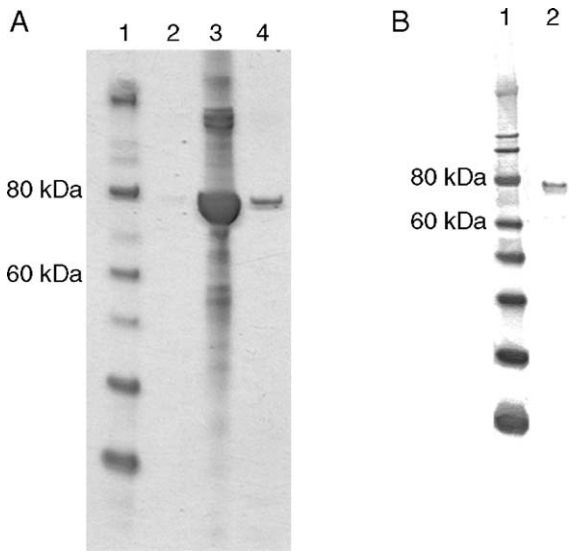

Fig. 4. Protein purification and Western blot analyses. A. Purification of His-tagged-AcRpoD. Lane 1, molecular weight size markers; Lane 2, no sample; Lane 3, Gu-HCl-solubilized inclusion bodies; Lane 4, purified refolded His-tagged-AcRpoD. The 80 and 60 kDa molecular weight bands are indicated on the left. B. Immunoblot of AcRpoD with 2G10 mAb directed toward RpoD. Lane 1, molecular weight size markers; Lane 2, purified refolded His-tagged-AcRpoD.

acrylic acid, results in the formation of the indole acrylic acid-*trp* repressor complex, which is unable to bind the *trp* operator, thereby permitting transcription of the endogenous *rpoD* (Morse et al., 1969). CAG20153 was transformed with pCF430AcRpoD, plated onto LBA+Cm+ indole acrylic acid, and transformants were picked and screened by PCR and double restriction digest with *Nhe* I and *Hin*D III to ensure that pCF430AcRpoD was stably maintained in CAG20153. CAG20153 with pCF430 served as a negative control. Two milliliters of LB+Cm+ indole acrylic acid were inoculated with CAG20153 pCFAcRpoD, CAG20153 pCF430 and CAG20153 and all three were incubated overnight at 37 °C on a rotary shaker. Cells were harvested, washed three times with LB and resuspended in 1 mL LB. Serial dilutions up to 10^{-7} of CAG20153 pCFAcRpoD, CAG20153 pCF430 and CAG20153 in LB were spotted onto LBA, LBA+ indole acrylic acid and LBA+ arabinose. Growth on all plates except LBA+ indole acrylic acid after 24 h incubation at 37 °C indicated cell viability and complementation of the mutant under growth-restrictive conditions.

3. Results and discussion

3.1. Cloning rpoD and sequence analyses

We cloned the *A. capsulatum rpoD* using the degenerate-inverse method depicted in Fig. 1 (GenBank accession no. DQ302476). Nucleotide sequence analyses indicate that the *A. capsulatum rpoD* has a putative ribosome binding site within 8 nucleotides of the GTG start codon. ORF Finder did not resolve any open reading frames, partial or complete, within 323 bp of the 5′-region of *rpoD*. Two hexanucleotide sequences, GTGAAG and TATAAG, separated by 18 nucleotides were observed within 40 nt upstream of the start codon thereby resembling −35 and −10 promoter elements. FindTerm identified a 29-nucleotide inverted repeat located 21 nucleotides downstream of the TAA stop codon, indicating a possible rho-independent termination stem-loop structure. BLASTX analyses of 615 bp downstream of the putative terminator stem-loop revealed 31% identity and 52% similarity to the C-terminal portion of a protein of unknown function, designated UPF0118, from *S. usitatus* Ellin6076. A 25-nt inverted repeat located 22 nt downstream from the TGA stop codon of this protein is also present, indicating that this ORF is translated, and likely transcribed, independently from *rpoD* and in the opposite direction (Fig. 2).

BLASTX analyses of the 1794-bp ORF indicated that it encodes a protein of 597 amino acids, with a molecular mass of 69.2 kDa and 57% identity and 75% similarity to the *S. usitatus* Ellin6076 major sigma factor protein. We compared the predicted amino acid sequence of *rpoD* to the major sigma factor protein sequences from *S. usitatus*, *E. coli*, *Bacillus subtilis* and *C. crescentus* by global alignment to identify sequence identities and structural motifs typical of bacterial major sigma factors. CLUSTALW analyses revealed that the *A. capsulatum* RpoD (AcRpoD) shared significant sequence identity with all major sigma factor protein sequences particularly in four highly conserved regions. Conserved domain searches highlighted subregions within *Ac*RpoD that have been functionally described: 1.1, autoinhibition of promoter binding (Camarero et al., 2002); 1.2, involved in core polymerase association and transcription initiation transitional states (Baldwin and Dombroski, 2001); 2.4/4.2, recognition of −10 and −35 promoter elements, respectively (Gross et al., 1998); and 3.0, recognition of the "extended-10" promoter element (Barne et al., 1997) (Fig. 3).

3.2. RpoD overexpression, purification and Western blot analysis

To further characterize the *A. capsulatum rpoD* gene product, we overexpressed the protein in an IPTG-inducible T7 polymerase-based system. SDS-PAGE analyses of purified AcRpoD

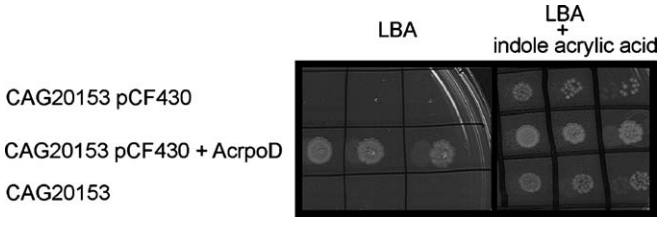

Fig. 5. Complementation of *E. coli* CAG20153 with *A. capsulatum* RpoD. Complementation of the CAG20153 strain harboring the chromosomal *rpoD* under the control of the *trp* promoter with the pCF430AcRpoD construct (middle row) containing the *A. capsulatum rpoD*. The CAG20153 pCF430, bearing no insert, (top row) and CAG20153 (bottom row) are controls. Indole acrylic acid induces expression from a *trp* promoter that drives expression of the endogenous *rpoD*, producing a growth permissive condition (right panel). In the absence of indole acrylic acid, cells require another source of active RpoD (left panel); therefore, growth of CAG20153 pCF430AcRpoD in the absence of indole acrylic acid indicates complementation. Absence of growth of CAG20153 and CAG20153 pCF430 strains on LBA indicates repression of expression of the endogenous *rpoD* by tryptophan in the media.

resolved a single band that migrates at 70 kDa (Fig. 4A, lane 4). A Western blot of refolded, His-tagged RpoD indicated a 70 kDa protein that was immunoreactive with RpoD-specific antibody (2G10) (Fig. 4B, lane 2). 2G10 confirmed the presence of a helix-turn-helix motif located in region 3.0 and 3.1 of the purified protein (Breyer et al., 1997). A second Western blot, using anti-polyhistidine mAb, confirmed the presence of the N-terminal hexahistidine fusion tag, which was exploited to purify over-expressed AcRpoD and distinguish it from contaminating host sigma factors (data not shown). These results indicate that *A. capsulatum rpoD* encodes a protein whose size and immunogenicity are consistent with those of other bacterial RpoD proteins.

3.3. Complementation of E. coli CAG20153

To determine whether the *A. capsulatum rpoD* gene product had an activity as a transcription factor, we introduced it into *E. coli* CAG20153, whose endogenous *rpoD* has been modified so that it is repressed by tryptophan. *A. capsulatum rpoD* restored growth to strain CAG20153 in LBA containing tryptophan and lacking indole acrylic acid, conditions that restrict growth of this strain (Fig. 5). The presence or absence of arabinose in the medium did not affect complementation. Therefore, *A. capsulatum rpoD* is expressed and can rescue strain CAG20153 under conditions when its endogenous *rpoD* is not expressed, indicating that the *A. capsulatum* RpoD is functional in *E. coli*.

4. Conclusion

As a first step toward understanding gene regulation mechanisms in the Acidobacteria, we cloned and characterized the *A. capsulatum* gene encoding the major sigma factor and characterized its gene product. The *A. capsulatum* RpoD was active in *E. coli* and its deduced amino acid sequence revealed structural motifs that are typical of bacterial major sigma factors, and the *A. capsulatum* RpoD cross-reacted with a monoclonal antibody specific for major sigma factors. Bacterial major sigma factors, including *A. capsulatum* RpoD, share considerable amino acid sequence conservation, which we exploited to amplify and clone *rpoD*. Since the *A. capsulatum* RpoD rescued CAG20153 under growth-restrictive conditions, it is likely that RpoD recognized the *E. coli* promoters required for growth. Interestingly, some major sigma factors are promiscuous in promoter element recognition and therefore facilitate cross-species gene expression, which suggests that the major sigma factor can play an important role in heterologous gene expression (Hakimi et al., 2000; Wosten, 1998). The work reported here provides a foundation for study of expression of genes from both cultured and not-yet-cultured Acidobacteria (Liles et al., 2003; Quaiser et al., 2003).

Acknowledgements

We thank Nancy Thompson for providing 2G10 mAb, Rachel Mooney for CAG20153, Jinwei Zhang for technical assistance, Robert Landick and Naomi Ward for very helpful discussions. This work was supported by the following grants: NIH Ruth L. Kirschstein NSRA Fellowship 5 F31 AI056675 to ZLS, NIH GM28575 to RRB, NSF DEB-0213048 to MRL and RMG, The NSF Microbial Observatories Program MCB-0132085 to JH, the Howard Hughes Medical Institute Professors Program to JH, and the University of Wisconsin-Madison. Disclaimer: RRB has a financial interest in NeoClone Biotechnology, Madison, WI, which markets the mAb 2G10.

References

Altschul, S.F., et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Ausubel, F.M.R., Brent, R.E., Kingston, D.D., Moore, J.G., Seidman, J.G., Smith, J.A., Struhl, K. (Eds.), 1994. Current Protocols in Molecular Biology. J. Wiley and Sons, New York, N.Y.

Baldwin, N.E., Dombroski, A.J., 2001. Isolation and characterization of mutations in region 1.2 of *Escherichia coli* sigma70. Mol. Microbiol. 42, 427–437.

Barne, K.A., Bown, J.A., Busby, S.J., Minchin, S.D., 1997. Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. EMBO J. 16, 4034–4040.

Barns, S.M., Takala, S.L., Kuske, C.R., 1999. Wide distribution and diversity of members of the bacterial kingdom Acidobacterium in the environment. Appl. Environ. Microbiol. 65, 1731–1737.

Bergendahl, V., Burgess, R.R., 2003. Studying sigma–core interactions in *Escherichia coli* RNA polymerase by electrophoretic shift assays and luminescence resonance energy transfer. Methods Enzymol. 370, 192–205.

Breyer, M.J., Thompson, N.E., Burgess, R.R., 1997. Identification of the epitope for a highly cross-reactive monoclonal antibody on the major sigma factor of bacterial RNA polymerase. J. Bacteriol. 179, 1404–1408.

Browning, D.F., Busby, S.J., 2004. The regulation of bacterial transcription initiation. Nat. Rev. Microbiol. 2, 57–65.

Burgess, R.R., Travers, A.A., Dunn, J.J., Bautz, E.K., 1969. Factor stimulating transcription by RNA polymerase. Nature 221, 43–46.

Camarero, J.A., et al., 2002. Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR. Proc. Natl. Acad. Sci. U. S. A. 99, 8536–8541.

Davis, K.E., Joseph, S.J., Janssen, P.H., 2005. Effects of growth medium, inoculum size, and incubation time on culturability and isolation of soil bacteria. Appl. Environ .Microbiol. 71, 826–834.

Gross, C.A., et al., 1998. The functional and regulatory roles of sigma factors in transcription. Cold Spring Harbor Symp. Quant. Biol. 63, 141–155.

Gruber, T.M., Bryant, D.A., 1997. Molecular systematic studies of eubacteria, using sigma70-type sigma factors of group 1 and group 2. J. Bacteriol. 179, 1734–1747.

Hakimi, M.A., Privat, I., Valay, J.G., Lerbs-Mache, S., 2000. Evolutionary conservation of C-terminal domains of primary sigma70-type transcription factors between plants and bacteria. J. Biol. Chem. 275, 9215–9221.

Joseph, S.J., Hugenholtz, P., Sangwan, P., Osborne, C.A., Janssen, P.H., 2003. Laboratory cultivation of widespread and previously uncultured soil bacteria. Appl. Environ. Microbiol. 69, 7210–7215.

Kishimoto, N., Kosako, Y., Tano, T., 1991. *Acidobacterium capsulatum* gen. nov., sp. nov.: an acidophilic chemoorganotrophic bacterium containing menaquinone from acidic mineral environment. Curr. Microbiol. 22, 1–7.

Liles, M.R., Manske, B.F., Bintrim, S.B., Handelsman, J., Goodman, R.M., 2003. A census of rRNA genes and linked genomic sequences within a soil metagenomic library. Appl. Environ. Microbiol. 69, 2684–2691.

Lonetto, M.A., Rhodius, V., Lamberg, K., Kiley, P., Busby, S., Gross, C., 1998. Identification of a contact site for different transcription activators in region 4 of the *Escherichia coli* RNA polymerase sigma70 subunit. J. Mol. Biol. 284, 1353–1365.

Marchler-Bauer, A., Bryant, S.H., 2004. CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 32, 327–331.

Morse, D.E., Mosteller, R.D., Yanofsky, C., 1969. Dynamics of synthesis, translation, and degradation of *trp* operon messenger RNA in *E. coli*. Cold Spring Harbor Symp. Quant. Biol. 34, 725–740.

Newman, J.R., Fuqua, C., 1999. Broad-host-range expression vectors that carry the L-arabinose-inducible *Escherichia coli araBAD* promoter and the *araC* regulator. Gene 227, 197–203.

Quaiser, A., et al., 2003. Acidobacteria form a coherent but highly diverse group within the bacterial domain: evidence from environmental genomics. Mol. Microbiol. 50, 563–575.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Wosten, M.M., 1998. Eubacterial sigma-factors. FEMS Microbiol. Rev. 22, 127–150.