**The Rhizobium etli σ\(^{70}\) (SigA) factor recognizes a lax consensus promoter**

Miguel A. Ramírez-Romero*, Irina Masulis, Miguel A. Cevallos, Víctor González and Guillermo Dávila

Programa de Genómica Evolutiva, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Apartado Postal 565-A, C.P 62210, Cuernavaca, Morelos, México

Received December 2, 2005; Revised and Accepted February 15, 2006

**ABSTRACT**

A collection of *Rhizobium etli* promoters was isolated from a genomic DNA library constructed in the promoter-trap vector pBMC535, by their ability to drive the expression of a *gusA* reporter gene. Thirty-seven clones were selected, and their transcriptional start-sites were determined. The upstream sequence of these 37 start-sites, and the sequences of seven previously identified promoters were compared. On the basis of the sequence conservation and mutational analysis, a consensus sequence CTTGACN\(^{16-23}\)TATNNT was obtained. In this consensus sequence, nine of twelve bases are identical to the canonical *Escherichia coli* σ\(^{70}\) promoter, however the *R.etli* promoters only contain 6.4 conserved bases on average. We show that the *R.etli* sigma factor SigA recognizes all *R.etli* promoters studied in this work, and that *E.coli* RpoD is incapable of recognizing them. The comparison of the predicted structure of SigA with the known structure of RpoD indicated that regions 2.4 and 4.2, responsible for promoter recognition, are different only by a single amino acid, whereas the region 1 of SigA contains 72 extra residues, suggesting that the differences contained in this region could be related to the lax promoter recognition of SigA.

**INTRODUCTION**

The *Escherichia coli* RNA polymerase holoenzyme consists of a core enzyme (σ25β\(^{′}\)), and one sigma factor (σ) subunit, which directs the core enzyme to specific sites on DNA to initiate transcription. In general, bacteria contain a main or housekeeping sigma factor, also known as sigma-70 (σ\(^{70}\)) which controls the expression of most of the genes, in association with the RNApol core (1-4). Sequence alignments of several σ\(^{70}\) factors from different bacteria reveal four conserved regions (1 to 4), that have been further divided into subregions, involved in different aspects of transcription initiation, including promoter recognition, promoter melting, initiation of RNA syntheses, abortive initiation and promoter escape (5).

The σ\(^{70}\) subunit of *E.coli* and *Bacillus subtilis* recognizes two types of promoter sequences: the first one, or ‘canonical promoter’, consisting of two hexamer sequences [TTGACA] and [TATAAT], centered respectively in the positions −10 and −35 relative to the transcription start-site (6,7). A helix-turn-helix motif located in the C-terminus of σ\(^{70}\) (subregion 4.2) recognizes the −35 box of the promoter, whereas a more centrally located α-helix (subregion 2.4) recognizes the −10 box (8). The second one, named ‘extended−10 promoter’ consists of an element [TATAAT] plus 2 nt [TG], located in the −15 and −14 positions. In this promoter type a −35 region or equivalent sequence does not exist. The N-terminal part of the subregion 3.0 (previously named 2.5) from σ\(^{70}\) is involved in the recognition of the conserved [TG] motif (9-11).

*Rhizobium etli* is a soil α-proteobacterium with the ability to colonize the roots of bean-plants to form nitrogen fixing nodules. The whole genomic sequence of *R.etli* CFN42 was recently concluded (12); GeneBank accession no CP000133, and consists of a circular chromosome and six large plasmids, with an average G + C content of 61.5%. *R.etli* contains a large number of sigma factors (one housekeeping σ\(^{70}\) gene, two copies of *rpoH*, two copies of *rpoN* and 18 genes of the extracistoplasmic factor group), a feature shared with other nitrogen fixing organisms, like *Bradyrhizobium japonicum*, *Mesorhizobium loti* and *Sinorhizobium meliloti* (4). The physiological roles and the mechanisms of promoter recognition by these sigma factors remain unknown. Moreover, in these organisms only a few well-characterized promoters are available.

*Correspondence may also be addressed to Guillermo Dávila. Tel: +52 777 329 16 90; Fax: +52 777 317 55 81; Email: davila@ccg.unam.mx

© The Author 2006. Published by Oxford University Press. All rights reserved.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oxfordjournals.org
The \( R.\text{etli} \) housekeeping \( \sigma^{70} \) gene (sigA) encodes a protein of 685 amino acid residues with a molecular weight of 77.18 kDa. The primary structure of SigA is very similar to RpoD (the \( \sigma^{70} \) factor of \( E.\text{coli} \)), especially in regions 2, 3 and 4, however SigA contains an extended region 1 (13). \( S.\text{meliloti} \), \( Caulobacter \text{sphaeroide} \) and \( Rhodobacter \text{sphaeroides} \) also have a larger SigA region 1, and most of their \( \sigma^{70} \) promoters that have been characterized are not transcribed by the \( E.\text{coli} \) RNAPol in vivo or in vitro, but in contrast, the RNAPol of these bacteria can initiate transcription of typical \( E.\text{coli} \) \( \sigma^{70} \) promoters (13–16). This observation suggests some differences between the transcriptional machineries of \( E.\text{coli} \) and the \( \alpha \)-proteobacteria, perhaps at the level of promoter recognition by the \( \sigma^{70} \) factor.

The knowledge of the promoter structures and their regulation in a particular organism is the first step towards the molecular understanding of the regulatory networks, an essential prerequisite to develop tools for genetic engineering.

To analyze the molecular basis of \( R.\text{etli} \) gene expression in this study, we identified, characterized and sequenced active promoters in \( R.\text{etli} \) under exponential growth conditions. A consensus promoter sequence was deduced, and possesses some differences to the promoter consensus recognized by \( \sigma^{70} \) in \( E.\text{coli} \) and \( B.\text{subtilis} \). We showed that the double helix stability and the stacking energy of the \( R.\text{etli} \) and \( E.\text{coli} \) promoter regions possess a minimal energy value around \(-10 \) box, suggesting that despite the differences in their promoter structures, the promoters of both bacteria preserve thermodynamic properties, which favor DNA denaturation and help the transcription initiation. We demonstrated that SigA specifically recognizes the \( R.\text{etli} \) promoters studied in this work, but in contrast, the \( E.\text{coli} \) RNAPol holoenzyme is incapable of recognizing them. Finally, we propose that \( R.\text{etli} \) contains a less strict \( \sigma^{70} \) that recognizes a lax consensus promoter structure.

**MATERIALS AND METHODS**

**Bacterial strains and growth conditions**

The bacterial strains and plasmids used in this work are listed in Table 1. \( E.\text{coli} \) strains were grown at \( 37^\circ \text{C} \) in Luria–Bertani (LB) medium. \( R.\text{etli} \) strains were grown at \( 30^\circ \text{C} \) in PY medium (17) or minimal medium containing 1.2 mM \( K_2\text{HPO}_4 \), 0.8 mM \( \text{MgSO}_4 \), 10 mM succinic acid, 10 mM \( \text{NH}_4\text{Cl} \), 1.5 mM \( \text{CaCl}_2 \) and 0.0005% \( \text{FeCl}_3 \), with the pH adjusted to 6.8 (18). Microaerobic conditions during growth were set up by putting the cultures in 150 ml bottles that were flushed with several volumes of a 1% oxygen-99% argon mixture and closed with an airtight stopper. Antibiotics were added at the following concentrations (in \( \mu\text{g} \cdot \text{ml}^{-1} \)): gentamicin 30; chloramphenicol 25; ampicillin 100; nalidixic acid 20. When it was used, X-Gluc (5-Br-4-Cl-3-indolyl-\( \beta \)-D-glucuronic acid, Research Organics) was added at a final concentration of 20 \( \mu\text{g} \cdot \text{ml}^{-1} \).

**DNA and RNA isolation and manipulation**

Genomic DNA was isolated using components and instructions from GenomicPrep cells and tissue DNA isolation kit (Amersham). Plasmid DNA was isolated as described by Sambrook et al. (19). DNA was restricted and ligated under the conditions specified by the enzyme manufacturer (Invitrogen). \( \text{Tag} \) DNA polymerase (Altaenzymes) was used for PCR. The PCR products were cloned using a pMOBlue blunt-ended kit (Amersham). RNA was isolated using components and instructions from the High pure RNA isolation kit (Roche).

**Table 1. Bacterial strains and plasmids used**

| Bacterial strains/plasmids | Relevant characteristics | Source/ references |
|---------------------------|--------------------------|------------------|
| **Strains**              |                          |                  |
| *Rhizobium etli*         |                          |                  |
| CFN42                    | Streptomycin resistant, nodulates *Phaseolus vulgaris* | (17)             |
| **Escherichia coli**     |                          |                  |
| DH5str                   | Host strain for plasmids | (47)             |
| BW21038                 | *gusA* derivative        | (48)             |
| DH10B                   | *lacZ* derivative        |                  |
| BL21(DE3)               | Host strain used for expression of target proteins | (22)             |
| **Plasmids**            |                          |                  |
| pBBR1MCS5               | Gentamycin resistant, cloning vector replicable in *R.\text{etli}* | (49)             |
| pBBMCS53                | *pBBR1MCS5* derivative carrying a promoter minus \( \beta \)-glucuronidase gene (*gusA*) | (20)             |
| pKK415                  | Tetraacyline resistant, derivative of pRK290 | (50)             |
| pLysS                   | Chloramphenicol resistant, production of lysozyme for repression of T7 polymerase | (22)             |
| pET-28a(+)              | Kanamycin resistant, expression vector pRK415 derivative carrying a 1726 bp fragment containing the *sigA* gene of the CFN42 strain | Novagen This work |
| pRKS70                  |                          |                  |
| pGUSphencC              | A pGUSphencC derivative in which the promoter element --10 was changed from [TACGAG] to [GCGGAG] | This work |
| pGUSphencC-10           | A pGUSphencC derivative in which the promoter element --35 was changed from [ATTGCG] to [AAGAGC] | This work |
| pGUSphencC-35           | A pGUSphencC derivative in which the promoter element --35 was changed from [ATTGCG] to [AAGAGC] | This work |
| pGUSphencC-1035         | A pGUSphencC derivative, but with mutated --10 and --35 elements from the promoter. Both elements were changed to the sequences indicated before | This work |

**Construction of a \( R.\text{etli} \) promoter library**

To isolate DNA fragments containing promoters, 10 mg chromosomal DNA of CFN42 strain was disrupted by nebulization during 2 min to 10 psi. Fragments ranging from 0.5 to 1 kb were purified from an agarose gel, blunted with Klenow and T4 pol enzymes, and ligated into the alkaline phosphatase derivative of pBBMCS53 derivative in which the promoter element --10 was changed from [TACGAG] to [GCGGAG]. This work.
Bacterial matings

pBBMCS53 derivatives were introduced by conjugation into \textit{R. etli} CFN42 using pRK2013 as a helper plasmid. Strains were grown in PY liquid medium to a stationary phase, mixed in a proportion of donor–helper–receptor of 1:1:2 on PY plates and incubated at 30°C overnight. Cells were resuspended in fresh PY medium, and serial dilutions were plated in PY nalidixic acid 20, gentamicin 30 and X-Glu. Clones containing fragments with a promoter were selected by blue/white screening of transconjugants.

\( \beta \)-Glucuronidase activity measurements

Overnight cultures of the strains carrying the desired constructs were grown in the presence of the appropriate selection. Cultures were diluted in fresh PY medium to an OD\( _{540} \) of 0.01 and grown to a final OD\( _{540} \) of 0.4. The culture (1 ml) was centrifuged and resuspended in a salt wash solution supplemented with chloramphenicol (100 mg/ml\(^{-1} \)). Quantitative \( \beta \)-glucuronidase assays were performed with \( p \)-nitrophenyl glucuronide substrate as described previously (21). Data were normalized to total-cell protein concentration by the Lowry method (19). The results presented in the figures are the mean of three independent experiments.

Transcription start-sites determination

Transcription start-sites were mapped by means of 5' RACE kit version 2.0 (Invitrogen). Briefly, DNA single chains (scDNA) were synthesized from 5 to 10 mg of total RNA using a gene-specific primer (GUS-LW5, CGATCCGACT-GAATGGCCAC) complementary to the region located in the position 96 to 117 from \textit{gusA}, and using a derivative of Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT). After first strand cDNA synthesis, the original mRNA template was removed by treatment with a RNase Mix (mixture of RNase H and RNase T1) and the scDNA was purified using a GlassMax DNA column. A homopolymeric tail was then added to the 3' end of the scDNA using Terminal deoxynucleotidyl transferase (TdT) and dCTP. In the next step, a PCR amplification was accomplished using \( Tau \) DNA polymerase, a second antisense primer (GUS-LW4, GTAAAC-TAAGGGACTGACCTGC) that was complementary to \textit{gusA} in the position 28 to 49, and a complementary homopolymer tail primer (AAP, GGCCACGCGTCGACTAGTACTATAGACGGGIGGGGIIAGGIIIG). When secondary bands were obtained, an additional PCR was done using an AAP primer and a third antisense primer (GUS-LW2, CTTTGGGCTATATCATGTCAT), that was complementary to the DNA region located in the position 1 to 21 from \textit{gusA}.

DNA sequencing

DNA sequencing reactions were performed using the GUS-LW primers and Big-Dye terminator kit version 3.1 in an automatic 310 DNA sequencer (Applied Biosystems).

Construction of promoter mutations

Mutant versions of the \textit{phemC} and \textit{lac} promoters were constructed by PCR using oligonucleotides carrying altered versions of the wild-type –35, –10 and both boxes simultaneously. Each PCR product was cloned in to pBBMCS53 vector and introduced into \textit{R. etli} CFN42 by triparental conjugation.

Overexpression and purification of the \textit{R. etli} SigA protein

A DNA fragment corresponding to the full-length \textit{sigA} gene of \textit{R. etli} was PCR amplified from total DNA, and cloned into pET-28a (+) (Novagen) to express \textit{sigA} in BL21 (DE3)plysS strain (22). The induction in exponential growth phase (OD\( _{540} = 0.4 \)) was obtained by the addition of isopropyl-\( \beta \)-thiogalactopyranoside (IPTG) up to the final concentration 1 mM. Cells were harvested after 4 h of cultivation and stored at –70°C. A total of 10 ml of culture usually yielded 2.5–3 mg of protein. Histidine tagged sigma protein was purified using HisTrap Kit (Amersham Biosciences). The cell pellet was suspended in 0.5 ml of binding buffer, containing 8 M urea and 0.25 mM phenylmethylsulfonyl fluoride (PMSF) [1 mM phosphate buffer (pH 7.4), 20 mM imidazole], and incubated 30 min at room temperature with gentle mixing. After sedimentation of cell debris, lysate was loaded on to Ni-column and equilibrated with a binding buffer. The desired protein was eluted by phosphate buffer, containing 400 mM imidazol, collected in 0.5 ml fractions, and analyzed in 10% SDS–PAGE. The applied procedure provides protein of 95% purity. To avoid the influence of six-aminoacide histidine tagg, His-Trap purified protein was treated by thrombin (Thrombin CleanCleave Kit, Sigma) after buffer exchange by dialysis. After the thrombin treatment protein was dialedyzed against TGED buffer [10 mM Tris–HCl (pH 7.6); 0.1 mM EDTA; 0.1 mM DTT; 0.1 M NaCl; 5% (v/v) glycerol] and storage buffer [10 mM Tris–HCl (pH 7.6); 0.1 mM EDTA; 0.1 mM DTT; 10 mM MgCl\(_2\); 0.2 M KCl; 50% (v/v) glycerol]. The final concentration of protein was evaluated by method of Bradford (Bio-Rad protein assay kit).

DNA binding shift assay

Linear DNA templates were randomly labeled with [\( \alpha \)-\( P ^{32} \)]dCTP in a standard PCR, purified from 3.5% PAGE and used for DNA–protein complex formation. RNA polymerase was assembled from core and \( \sigma \)^70 protein as it was described by Tang \textit{et al.} (23). Binary complexes were allowed to preform for 20 min at 30°C, challenged by heparin (200 mg/ml) and immediately loaded on to 3.5% native PAGE, prewarmed up to the same temperature. After electrophoresis, gels were dried and radioautographed.

Software tools

Multiple DNA alignments were done using the programs CLUSTALW version 1.8.1 (24) and MultiAlin version 5.4.1 (25). Alignments were visually inspected and adapted with respect to the known transcriptional start-sites. Consensus matrices were generated and selected using WCONSENSUS version v5c (26). Briefly, to identify the –10 box, the first 20 bp upstream from the transcription start-site (+1) were used. To identify the –35 box, the 20 bp upstream to the –20 position were used as an input for WCONSENSUS. Algorithms available from server http://www.icgeb.trieste.it/dna (27) were used to evaluate double helix stability (28) and stacking energy (29). Calculations for each parameter were carried out with a window size of 10 and 20 nt, providing...
similar results. For each promoter position, the average value was calculated for the whole set of analyzed sequences.

RESULTS

Isolation of *R. etli* promoter sequences

To determine the general structure of the *R. etli* promoters, a random genomic library with an average insert size of 0.5 kb was constructed in the promoter-test vector pBBMCS53 and introduced by conjugation into *R. etli* CFN42. Direct screening on PY plates containing X-Gluc allowed the selection of transconjugants with an active reporter gene (*gusA*), suggesting that the constructs contained *R. etli* promoter regions. The β-glucuronidase activity of 100 transconjugants was determined in PY liquid cultures in exponential and stationary phases of growth. Sixty-eight of these clones showed β-glucuronidase activity under both conditions, while the remaining 32 transconjugants expressed the reporter gene only in PY plates, and for this reason were not used in further experiments. The ability to express the reporter gene of the 68 strains described above was tested in different conditions: low oxygen (1%), minimal medium (MM) and at different temperatures (30, 42°C). All but three expressed the reporter gene in all conditions, but at different levels, suggesting that the transconjugants carry constructs with constitutive promoters. The DNA sequences of 36 inserts of these constructs were determined. These promoter sequences, plus seven *R. etli* promoters experimentally identified in other studies were selected for subsequent analyses (30–35).

Thirty-three promoters were mapped within intergenic regions; eight within the coding regions of the immediately preceding gene, and three in DNA sections that contained overlapped coding regions. The genetic compartments of the 44 promoters of our collection were next: thirty-nine were located on the chromosome, two on plasmid p42d and one in each of these plasmids: p42b, p42e and p42f.

Identification of transcriptional start-sites

The transcription start-sites are good reference points to identify the upstream conserved sequence characteristics of the promoters. Thus, the 5' ends of the mRNAs synthesized from each of the selected 36 promoter regions selected, were successfully identified by 5' RACE using the total RNA isolated from PY cultures in exponential growth conditions of the transconjugants. In general, a single RT–PCR product was obtained, but in one clone, two conspicuous amplification products were observed. In this case the bands were independently isolated, cloned and sequenced, showing that these products are transcripts associated to the same gene.

Thirty-eight of the transcription start-sites (86%) fell within the first 250 bp upstream from coding regions, and four transcription start-sites corresponded to the nucleotide A of the start codon ATG from their respective ORF (Table 2).

| Gene          | Distance between +1 to ATG (bp) | Gene product                        | References |
|---------------|---------------------------------|-------------------------------------|------------|
| cRNA          | NA                              | Antisense RNA regulator             | (30)       |
| repA          | 45                              | Plasmid partitioning protein RepAd  | (31)       |
| thiC          | 209                             | Thiamine biosynthesis               | (32)       |
| recA          | 37                              | Recombinase A                       | (33)       |
| lipA          | 63                              | Lipoic acid synthetase              | (34)       |
| nifB          | 71                              | Nitrogen regulation                 | (35)       |
| nifR          | 28                              | Nitrogen regulation                 | (35)       |
| sigA          | 230                             | RNA polymerase sigma factor         | This work  |
| ypyf00079     | 18                              | Probable methyl-accepting           | This work  |
| Glycomannan protein |                     | chemotaxis protein                | This work  |
| glyQ          | 27                              | Glycolyl–RNA synthetase alpha       | This work  |
|                |                                 | subunit                            |            |
| yych00468     | 37                              | Hypothetical conserved protein      | This work  |
| yych00334     | 54                              | Hypothetical conserved protein      | This work  |
| yych01154     | 229                             | Hypothetical conserved protein      | This work  |
| Eda           | 35                              | KRG-KDPG aldolase                   | This work  |
| yych00528     | 19                              | Hypothetical conserved protein      | This work  |
| enbB          | 29                              | GTP-binding protein                 | This work  |
| 16SrRNA       | 67                              | 16S ribosomal RNA                   | This work  |
| yych01141     | 67                              | Hypothetical conserved protein      | This work  |
| yych00182     | 23                              | Hypothetical conserved protein      | This work  |
| yych00739     | A                               | Hypothetical conserved protein      | This work  |
| hemC          | 18                              | Porphobilinogen deaminase           | This work  |
| clpF          | 55                              | ATP-dependent Clp protease          | This work  |
| rpoH1         | 101                             | Proteolytic subunit 2               | This work  |
| rpoH1         | 101                             | RNA polymerase sigma factor protein | This work  |
| arsG          | 164                             | Phospho-2-dehydro-3-                | This work  |
| yych00196     | 41                              | deoxyheptonate aldolase             | This work  |
| ureA          | 34                              | Putative transcriptional regulator  | This work  |
|                  |                                 | Urease (urea amidohydrolase)        | This work  |
| yych00131     | 268                             | Putative transcriptional regulator  | This work  |
| ndhB          | 24                              | Probable transcriptional regulator  | This work  |
|                  |                                 | Urease (urea amidohydrolase)        | This work  |
| yych00947     | 29                              | Putative tail-specific protein       | This work  |
| yych00197     | 21                              | Putative tail-specific protein       | This work  |
| yych00168     | 21                              | Hypothetical conserved protein      | This work  |
| yych00468     | 18                              | Hypothetical conserved protein      | This work  |
| yych00187     | 18                              | Hypothetical conserved protein      | This work  |
| yych00532     | 18                              | Hypothetical conserved protein      | This work  |
| yych00196     | 18                              | Hypothetical conserved protein      | This work  |
| yych00532     | 18                              | Hypothetical conserved protein      | This work  |
| ynych00076    | 18                              | Hypothetical conserved protein      | This work  |
| exor2         | 242                             | Exopolysaccharide production        | This work  |
|                |                                 | negative regulator                  |            |

*Table 2. Distances between the experimentally determined transcriptional start-sites and the ATG for each gene*

NA: not applicable

*pseudogene; A, adenine from ATG; NA, non apply.

The genetic compartments of the 44 promoters of our collection were next: thirty-nine were located on the chromosome, two on plasmid p42d and one in each of these plasmids: p42b, p42e and p42f.

Determination of a promoter consensus sequence

The promoter consensus sequence was derived by aligning 50 bp upstream from the transcription start-site of each promoter region. The identification of −10 and −35 boxes was corroborated with the WCONSENSUS program as described in Materials and Methods. As shown in Figure 1a, over-represented nucleotides clearly appeared around the −35 region and, in a lesser degree, in the −10 region. Then, the putative −10 consensus box [TATNNNT] showed four identical positions to the −10 box of the *E.coli* consensus promoter [TATAAT] recognized by σ70 factor. The 3' end of the
The -10 extended promoter. The -35 box consensus sequence [CTTGAC], is identical in five positions to the equivalent box of the E.coli consensus promoter [TTGACA] recognized by σ70. The spacing between the -35 and -10 regions ranges from 16 to 23 bp.

Figure 1. (a) Alignment of the promoter regions identified in *R.etli*. The position of the -10 and -35 elements are indicated at the top of the alignment. The -10 and -35 elements identified by other authors are indicated within boxes. The nucleotides that appear in >51% of the cases in any given position are marked in yellow. The transcription start-sites are marked with +1. Alternative promoters with suboptimal properties are underline. The consensus sequence derived from the alignment are indicated at the bottom of the alignment. The promoters identified were: 1, prepA; 2, pctRNA; 3, precA; 4, plipA; 5, pthiC; 6, pnifR promoter 1; 7, pnifR promoter 2; 8, ppsA; 9, pypf00079; 10, ppyG; 11, pyhch00468; 12, pyhch00334; 13, pyhch01154; 14, peda; 15, pyhch00528; 16, pengB; 17, pl65 rI RNA; 18, pyhch0141; 19, pyhch00226; 20, pyhch00739 pseudogene; 21, phmcC; 22, pcplP; 23, prpsH1; 24, paroG; 25, pyhch00196; 26, pureA; 27, pyhch00131; 28, pndB; 29, pype00047 pseudogene; 30, pyhch00197; 31, ppygA; 32, pcocH; 33, pyhch00552; 34, pyhch00096; 35, pyhch0132; 36, prpsN promoter 1; 37, prpsN promoter 2; 38, pyhch00326; 39, prpsP; 40, pyhch01115; 41, pyhch00280; 42, pisP. (b) Distribution of spacer lengths for the 44 promoters of the (a) (c) Distribution of the distance between the 3' of the -10 element (TATNT) and the +1.

*R.etli* -10 consensus boxes was positioned from 3 to 10 bp upstream from the transcriptional start-site (Figure 1c). Six promoter regions (14%) showed the dinucleotide [TG] in positions -15 and -14, which resembled the structure of the E.coli -10 extended promoter.
from 16 to 23 bp (Figure 1b). The promoter sequences contain 6.4 conserved bases on average, of the 12 conforming −10 and −35.

Suboptimal promoters were identified in 11 of the 44 promoter regions as shown in Figure 1a, but they were not taken in account to generate the consensus promoter sequence for the following reasons: none of them contained both hexameric boxes recognized by the best statistic parameters of WCONSENSUS. Nevertheless, in six of these promoters the −35 box overlapped with one proposed by us. In one case the −10 box overlapped the region proposed here. Moreover, some of the suboptimal promoters contained the spacer region too long (up to 27 bp) or to short (11 or 12 bp). In two cases the transcription start-sites were located far of the −10 boxes (up to 20 bp).

The results presented here showed that the R.etli promoter consensus possesses some similarities with the E.coli promoter consensus recognized by the σ70 factor, but the variability in the nucleotide frequency in each position of the consensus was greater than that observed in E.coli promoters (7).

Thermodynamic properties of the R.etli promoter regions

It was recently demonstrated in several bacteria genomes, that the intergenic regions containing promoters are generally less stable and less flexible than coding regions (36). To determine if these thermodynamic properties are also present in the promoter regions of R.etli, the stacking energy and the double helix stability were evaluated in our collection of 44 promoters. For this analysis the sequence of 200 bp upstream and 100 bp downstream from transcription start-site of all members of the promoter collection were taken in account. As an internal control, a set similar in size and length was constructed with the immediately downstream coding region of each member of the promoter collection. The same properties were predicted and contrasted with two E.coli equivalent sets: one containing 44 randomly chosen promoter sequences experimentally characterized and putatively recognized by σ70, and the other of 44 coding region sequences that corresponded to the genes controlled by the selected promoters. As shown in Figure 2a, the distribution profiles of double helix stability for coding and no coding regions of R.etli were similar, however, a significative local minimum around the −10 element was detected in our collection of promoter regions. The average stability of the E.coli coding regions was higher than the average stability of their promoter regions, and like R.etli, a prominent feature of the promoter regions was a minimum stability around the −10 box (Figure 2b).

The stacking energy profiles of R.etli and E.coli promoter regions were variable, but with a tendency to low negative values (low stability), nevertheless local minimum values were located around the −10 box. In contrast, the stacking energy profiles of R.etli and E.coli coding regions were similar: both showed more negative values that corresponded to great stability (Figure 2a and b). These results suggest that despite the variability in the nucleotide composition of R.etli promoters, these regions possess thermodynamic and structural properties similar to the E.coli promoter regions.

Promoter sequence mutagenesis

To demonstrate that the promoter consensus proposed here is correct, the DNA sequence of three R.etli mutant promoters (from repA, the antisense RNA gene present in the p42d repABC operon and recA) with the reduced transcriptional activity previously characterized was located in our alignment. As shown in Figure 3, these mutations fall in the regions centred in −10 and −35 of the consensus, suggesting that these elements were correctly identified.

Additionally the promoter consensus was verified in the following way: the constitutive promoter (phemC) of hemC, a gene encoding a porphobilinogen deaminase protein, was selected for a mutagenic analysis. This promoter possessed five conserved positions from consensus and showed moderate activity. Three mutant derivatives were constructed that deviated the promoter sequence from the consensus. First, the putative −35 element was altered, changing sequence [ATTGCC] to [AAGACG] to generate the construct pGU-SphemC-35. Second, the putative −10 element was changed from [TACGAG] to [GGCGAG] to produce construct pGU-SphemC-10. Third, the putative wild-type −35 and −10 elements were changed at the same time to the sequences indicated before, to generate plasmid pGUSphemC-1035. These constructs were introduced into R.etli CFN42, and the β-glucuronidase activities of the transconjugants were determined. As shown in Figure 3, all mutant constructs showed a low level of β-glucuronidase activity, indicating that the −35 and −10 elements of the R.etli promoter consensus were correctly identified.

Activity of the R.etli promoters in E.coli

As described in the previous section, the R.etli σ70 promoter consensus sequence is more variable than that described for E.coli. This observation suggests that the E.coli σ70 is unable to recognize the R.etli promoters or recognizes them less efficiently. To prove this possibility, the 36 transcriptional fusions of the collection were introduced into the BW21038 (gusA) strain, and their β-glucuronidase activities were evaluated in plates supplemented with X-Gluc. In this experiment only three promoters were functional, but showed low β-glucuronidase activity. These results suggest that the 33 transcriptional fusions are poorly or unrecognized by the E.coli sigma factors. It is possible than the E.coli σ70 is incapable of binding the R.etli promoters because it is more stringent in its recognition pattern, and imposes a strict functional barrier. If this hypothesis is true, then the R.etli σ70 (SigA) activates the transcriptional fusions described above in E.coli. To demonstrate this, all strains containing the gusA fusions were transformed with a plasmid expressing the R.etli σ70 gene under the control of the lac promoter and the β-glucuronidase activity of the transformants were evaluated in LB plates containing X-Gluc. Representative results are shown in Figure 4. The expression of sigA allowed the gusA fusions to express their reporter gene in the 36 R.etli promoter regions studied here, however, the strains containing sigA constructs were unstable, and a mixture of blue and white colonies was observed in the plates (Figure 4). The results suggest that R.etli σ70 is responsible for the transcriptional activity of these promoters, and that the structural variability of R.etli promoters cannot be recognized by E.coli σ70.
The \textit{E.coli} \(\sigma^{70}\) is incapable of binding \textit{in vitro} \textit{R.etli} promoters

To verify if the \textit{R.etli} promoters are recognized by the \textit{E.coli} \(\sigma^{70}\) factor, electrophoretic mobility shift assay (EMSA) experiments were done using \textit{E.coli} RNApol containing \textit{E.coli} \(\sigma^{70}\) and a PCR product containing the promoter of \textit{hemC} gene (\textit{phemC}), plus the PCR products of four promoter regions selected at random: \textit{ureA}, \textit{cycH}, \textit{yhch00197} and \textit{yhch00468}. As a negative control, a PCR product of a mutant
derivative of promoter p\(\text{hemC}\) containing the changes in the 
\(-10\) and \(-35\) elements described previously was used, and as
a positive control, a PCR product containing the \(\text{lac}\) promoter
was used (\(\text{lac}\) promoter is functional in \(\text{R.etli}\)). Representative
results are shown in Figure 5a. In all experiments the \(\text{E.coli}\) RNApol holoenzyme was unable to bind the \(\text{R.etli}\) promoters,
indicating that \(\text{E.coli}\) \(\sigma^{70}\) cannot recognize \(\text{R.etli}\) promoters.

**SigA binds in vitro \(\text{R.etli}\) promoters**

To determine if SigA is capable of recognizing the \(\text{R.etli}\) promoters, binding shift experiments were done using a chimeric
RNApol holoenzyme composed of \(\text{E.coli}\) RNApol core and
\(\text{R.etli}\) \(\sigma^{70}\) factor assembled in vitro. As a target DNA,
the PCR products containing the promoters mentioned
above (\(\text{p}\text{hemC}\) and the promoters of \(\text{ure}\), \(\text{cycH}\), \(\text{yhch00197}\) and \(\text{yhch00468}\) genes) were used in these experiments.
Representative results are shown in Figure 5. The chimeric
RNApol holoenzyme bound the five PCR products, but at low
efficiency (Figure 5b). In contrast, no DNA binding shifts were
observed using a PCR control with the mutant promoter
\(\text{p}\text{hemC}\) (Figure 5c). These results showed that the chimeric
RNApol holoenzyme specifically recognizes the \(\text{R.etli}\) promoters.

**Comparative analysis of the \(\sigma^{70}\) consensus sequences
from different bacteria**

The \(\text{R.etli}\) \(\sigma^{70}\) promoter consensus was compared with those of
the: (i) \(\alpha\)-proteobacteria \(\text{C.crescentus}\) (G + C\% of 67) (37)
and \(\text{Rhodobacter capsulatus}\) (G + C\% of 67) (38); (ii) the

---

**Table:**

| Sequence | Promoter | % residual activity | Reference |
|----------|----------|---------------------|-----------|
| Wild-type | CTGCCGAGAGAATCCGATCAC | 30 |
| Mutant 1  | CTGCCGAGAGAATCCGATCAC | 1 |
| Mutant 2  | CTGCCGAGAGAATCCGATCAC | 0 |
| Mutant 3  | CTGCCGAGAGAATCCGATCAC | 5 |
| Mutant 4  | CTGCCGAGAGAATCCGATCAC | 1 |
| Mutant 5  | CTGCCGAGAGAATCCGATCAC | 100 |
| Mutant 6  | GATGCCGAGAGAATCCGATCAC | ND | 33 |
| Mutant 7  | GATGCCGAGAGAATCCGATCAC | 47 |
| Mutant 8  | GATGCCGAGAGAATCCGATCAC | 18 |
| Mutant 9  | GATGCCGAGAGAATCCGATCAC | 10 |
| Consensus | CTGGAGA | PATINT |

---

**Figure 3.** Genetic analysis of \(\text{R.etli}\) promoter regions All \(\text{R.etli}\) mutant promoters with reduced transcriptional activity previously characterized, plus mutant
derivatives of the \(\text{p}\text{hemC}\) promoter constructed in this work were aligned with the \(\text{R.etli}\) promoter consensus. The changes that produced a reduced transcriptional
activity are marked with lower-case letters.

**Figure 4.** \(\beta\)-glucuronidase activities of the \(\text{R.etli gus}\)\(_{A}\) fusions in \(\text{E.coli}\) Qualitative evaluation of the \(\beta\)-glucuronidase activities of \(\text{E.coli}\) BW21038 transformants containing different \(\text{R.etli}\) promoter regions fused to the \(\text{gus}\) gene of the promoter-trap vector pBBMCS53, with or without the expression of \(\text{sigA}\) present in the
plasmid pRK70. Each plate included positive (pHER; + symbol) and negative (pBBMCS53; – symbol) controls. Each pair of strains, marked with the same
number, contain the same transcriptional fusion. Strains marked with an asterisk harbor the pRK70, which expresses the \(\text{sigA}\) gene of \(\text{R.etli}\). The strains contain the
following promoter regions: 1, pglyQ; 2, pyhch1154; 3, prpoHI; 4, pyhch0528; 5, pyych00326; 6, pyych00096; 7, pyych00468; 8, pyych00334 and 9, peda.
DISCUSSION

*R.etli* is a bacterium that inhabits diverse environments and has genome rich in genes related to transcriptional regulation, suggesting that it contains a complex genetic regulation system. To characterize the molecular basis of the transcription, and particularly, what kind of sequences act as promoters in *R.etli*, we identified and characterized a promoter consensus sequence recognized by the *R.etli* $\sigma^{70}$ factor (SigA) in this paper.

The *R.etli* $\sigma^{70}$ promoter consensus sequence [CTTGAC] N$_{16–23}$ [TATNNT] was deduced from 44 promoter regions. Using site-directed mutagenesis of one representative *R.etli* promoter isolated in this work, and a collection of mutants of three promoter regions described previously, we found that all changes fell within the $-10$ or $-35$ boxes identified in this work, suggesting that the consensus sequence is correct. Moreover, two different lines of evidence indicate that this promoter consensus sequence is recognized by the *R.etli* $\sigma^{70}$. The first one is that gusA fusions containing *R.etli* promoters are not functional in *E.coli*, but all regain their activity when complemented with SigA. Second, SigA is able to bind to the *R.etli* promoters but not to their mutant derivatives that partially eliminate the $-35$ or $-10$ boxes.

The location of the 44 promoter regions in the *R.etli* genome showed that 33 are located within intergenic regions, and that 8 (18%), fall within coding regions. This situation has also been observed in other bacteria, i.e. in *E.coli* 18% of the promoters are within genes (41), and around of 20% in *Prochlorococcus marinus* (42).

Of the promoters 86% analyzed here were located less than 250 bp from the translation start codons of their genes. This property has been widely studied in *E.coli*, where it is known that 90% of the promoters are located less than 250 bp upstream from the gene that they regulate (41).

Although the *R.etli*, *B.subtilis* and *E.coli* promoter consensus display a common structure, the *R.etli* consensus sequence contains only 6.4 conserved bases on average, revealing a more variable composition than in the *E.coli* and *B.subtilis* consensus, where 7.9 and 9.1 nt are conserved on average, respectively (7,39,40). Additionally, the *R.etli* promoter consensus shows that the sequence length between the $-10$ and $-35$ boxes varies from 16 to 23 nt, nevertheless, 25 of the promoters (36.81%) present a distance of 17 ± 1 nt.
This variation range is similar to that found in E.coli (15 to 21 nt) and in B.subtilis (14 to 20 nt). The percentage of promoters with $17 \pm 1$ nt in length between the hexameric boxes, is quite low compared to E.coli (92%) and B.subtilis (84.5%) (6,39). The distance between +1 nt and −10 box 3’ end fluctuates from 3 to 10 bases, as found in E.coli and B.subtilis (6,39). Six promoters (14%) were found with a putative −10 extended sequence, and their frequency is lower than those reported for this kind of promoter in E.coli (20%) and B.subtilis (45%) (39,43,44).

A computer analysis of DNA double helix stability and stacking energy showed that the promoter regions presented a minimum value for both parameters around −10 box, as found in E.coli promoters. Having these results in mind, it is reasonable to conclude that in spite of their sequence heterogeneity, the promoter regions preserve thermodynamic properties which have been selected to favor DNA denaturation and help to start the transcription. Similar conclusions have been drawn for other bacterial promoter regions (36,45).

The R.etli promoters tested are not recognized by E.coli σ70, however, the E.coli lac promoter is recognized by R.etli σ70, suggesting that R.etli σ70 is more promiscuous than E.coli σ70.

The deduced primary structure of SigA is 47% similar to the E.coli σ70, especially in regions 2.4 and 4.2, which are responsible for promoter recognition. The most notable difference between SigA and RpoD is the larger size of region 1 of SigA, which contains 72 extra amino acids. This region could be related to the lac promoter recognition of SigA. Thus, E. coli σ70 being the strictest factor for promoter recognition, it allows less structural variation, which is reflected in a robust consensus sequence. The opposite case is observed in α-proteobacteria, where the promiscuous σ70 allows a larger variation in its promoter structure.

A promiscuous σ70 factor could be an adaptive advantage for bacteria like the α-proteobacteria that inhabit environments with high biological and environmental diversity, because they can take the benefits of the genetic information acquired by horizontal transfer in a fast and efficient way, avoiding a barrier that could limit the acquisition of new functions required for the adaptation to new environmental conditions.

Our comparison of the consensus of several bacteria suggests that differences in the genomic G + C content are related to promoter structure: the bacteria with high G + C content tend to be more variable, and to have −10 boxes different in sequence to those from bacteria with low G + C content, and vice versa. Previous evidence in other organisms supports this conclusion (46).

**ACKNOWLEDGEMENTS**

The authors wish to thank Marco Zúñiga, Rosa Elena Gómez Barreto, Ismael Hernández, Patricia Bustos, Rosa Isela Santamaría and José Espíritu for their skillful technical support. The authors also thank Paul Gaytán and Eugenio López for the synthesis of oligonucleotides. This work was funded by CONACYT grants N-028, N-U46333-Q for emergent areas, and the Universidad Nacional Autónoma de México. Funding to pay the Open Access publication charges for this article was provided by Universidad Nacional Autónoma de México.

Conflict of interest statement. None declared.

**REFERENCES**

1. Lonetto,M., Gribskov,M. and Gross,C.A. (1992) The σ70 family: sequence conservation and evolutionary relationships. J. Bacteriol., 174, 3843–3849.
2. Wosten,M.M.S.M. (1998) Eubacterial sigma-factors. FEMS Microbiol. Rev., 22, 127–150.
3. Maeda,H., Fujita,N. and Ishihama,A. (2000) Competition among seven Escherichia coli sigma subunits: relative binding affinities to the core RNA polymerase. Nucleic Acids Res., 28, 3497–3503.
4. Mittenhuber,G. (2002) An inventory of genes encoding RNA polymerase sigma factors in 31 completely sequenced eubacterial genomes. J. Mol. Microbiol. Biotechnol., 4, 77–91.
5. Paget,M.S. and Helmann,J.D. (2003) The σ70 family of sigma factors. Genome Biol., 4, 203.1–203.6.
6. Harley,C.B. and Reynolds,R.P. (1987) Analysis of E.coli promoter sequences. J. Mol. Biol., 15, 2343–2361.
7. Lisser,S. and Margalit,H. (1993) Compilation of E.coli mRNA promoter sequences. Nucleic Acids Res., 21, 1507–1516.
8. Murakami,K.S., Masuda S. and Darst,S.A. (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. Science, 296, 1280–1284.
9. Barne,K.A., Bown,J., Bushy,S.J. and 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.
10. Campbell,E.A., Muzzin,O., Chlenov,M., Sun,J.L., Olson,C.A., Weinman,O., Trester-Zedlitz,M.L. and Darst,S.A. (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. Mol. Cell., 9, 527–539.
11. González et al. (2005) The partitioned Rhizobium etli genome: Genetic and metabolic redundancy in seven interacting replicons. Proc. Natl Acad. Sci. USA, 103, 3834–3839.
12. Luka,S., Patriarca,E.J., Riccio,A., Iaccarino,M. and Defez,R. (1996) Cloning of the rpoD analog from Rhizobium etli: sigA of R.etli is growth phase regulated. J. Bacteriol., 178, 7138–7143.
14. Peck,M.C., Gaal,T., Fisher,R.F., Gourse,R.L. and Long,S.R. (2002) The RNA polymerase alpha subunit from Shinnorizobium mellotli can assemble with RNA polymerase subunits from Escherichia coli and function in basal and activated transcription both in vitro and in vitro. J. Bacteriol., 184, 3808–3814.
15. Cullen,P.J., Kaufman,C.K., Bowman,W.C. and Kranz,R.G. (1997) Characterization of the Rhodobacter capsulatus housekeeping RNA polymerase. In vitro transcription of photosynthesis and other genes. J. Biol. Chem., 272, 27266–27273.
16. Wu,J., Ohta,N., Benson,A.K., Ninfa,A.J. and Newton,A. (1997) Purification, characterization, and reconstitution of DNA-dependent RNA polymerases from Caulobacter crescentus. J. Biol. Chem., 272, 21558–21564.
17. Noel,K.D., Sánchez-A., Fernández-Leemans,J. and Cevallos,M.A. (1984) Rhizobium phaseoli symbiotic mutants with transposon Tn5 insertions. J. Bacteriol., 158, 148–155.
18. Bravo,A. and Mora,J. (1988) Ammonium assimilation in Rhizobium phaseoli by the glutamine synthetase-glutamate synthase pathway. J. Bacteriol., 170, 980–984.
19. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular Cloning. A Laboratory Manual. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
20. Corvera,A., Prom,E., Prom,J.C., Martinez-Romero,E. and Romero,D. (1999) The nolL gene from Rhizobium etli determines nodulation efficiency by modulating the acetylation of the fucosyl residue in the nodulation factor. Mol. Plant-Microbe Interact., 12, 236–246.
21. Wilson,K.J., Hughes,S.G. and Jefferson,R.A. (1992) The Escherichia coli gus operon, induction and expression of the gus operon in E.coli and the occurrence and use of GUS in other bacteria. In Gallagher,S.R. (ed.), GUS Protocols, Using the Gus Gene as a Reporter of Gene Expression. Meth. Enzymol., 185, 60–89.
22. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular Cloning. A Laboratory Manual. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
23. Tang,H., Severinov,K., Goldfarb,A. and Eibright,R.H. (1995) Rapid RNA polymerase genetics: one-day, no column preparation of reconstituted recombinant Escherichia coli RNA polymerase. Proc. Natl Acad. Sci. USA, 92, 4902–4906.
24. Higgins,D., Thompson,J., Gibson,Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, optimization potential function for the calculation of nucleic acid interaction energy I. Base stacking. Biopolymers, 15, 563–577.
25. Vlachovicek,K., Kajdi,L. and Pongor,S. (2003) DNA analysis servers: plot.it, bend.it, model.it and IS. Nucleic Acids Res., 31, 3686–3687.
26. Sugimoto,N., Nakano,S., Yoneyama,M. and Honda,K. (1996) Improved thermodynamic parameters and helix initiation factors to predict stability of DNA duplexes. Nucleic Acids Res., 24, 4501–4505.
27. Goodfried,R.L., Reis,R., Brown,D.L. and MacElroy,R.D. (1978) An optimized potential function for the calculation of nucleic acid interaction energies I. Base stacking. Biopolymers, 17, 2341–2360.
28. Venkova-Canova,T., Seberov,N.E., Ramírez-Romero,M.A. and Cevallos,M.A. (2004) Two discrete elements are required for the replication of repABC plasmid: an antisense RNA and a stem-loop structure. Mol. Microbiol., 54, 1431–1444.
29. Ramírez-Romero,M.A., Téllez-Sosa,J., Barrios,H., Pérez-Osgeuera,A., Rosas,V. and Cevallos,M.A. (2001) RepA negatively autoregulates the transcription of the repABC operon of the Rhizobium etli symbiotic plasmid basic replicon. Mol. Microbiol., 42, 195–204.
30. Miranda-Rios,J., Moreira,C., Taboada,H., Dávados,A., Encarnación,S., Mora,J. and Sbor, M. (1997) Expression of thiamine biosynthetic genes (thiCOGE) and production of symbiotic terminal oxidase cbb3 in Rhizobium etli. J. Bacteriol., 179, 6887–6893.
31. Tapia,A., Fernández-de Henestrosa,A.R. and Barbé,J. (1997) Characterization of the promoter of the Rhizobium etli recA gene. J. Bacteriol., 179, 1573–1579.
32. Tait,R., Riccio,A., Iaccarino,M. and Patriarca,E.J. (1997) Cloning and transcriptional analysis of the lps (lipoic acid synthetase) gene from Rhizobium etli. FEMS Microbiol. Lett., 149, 165–172.
33. Patriarca,E.I., Riccio,A., Tait,R., Colonna-Romano,S., Iaccarino,M. and Delez,R. (1993) The ntrBC genes of Rhizobium leguminosarum are part of a complex operon subject to negative regulation. Mol Microbiol., 9, 569–577.
34. Peeters,A.G., Jensen,L., Brunak,S., Staerfeldt,H.H. and Ussery,D.W. (2000) A DNA structural atlas for Escherichia coli. J. Mol. Biol., 299, 907–930.
35. Malakooti,J., Wang,S.P. and Ely,B. (1995) A consensus promoter sequence for Caulobacter crescentus genes involved in biosynthetic and housekeeping functions. J. Bacteriol., 177, 4372–4376.
36. Richard,C.L., Tandon,A. and Kranz,R.G. (2004) Rhodobacter capsulatus nfa1 promoter: High-GC–10 regions in high-GC bacteria and basis for their transcription. J. Bacteriol., 186, 740–749.
37. Helman,J.D. (1995) Compilation and analysis of Bacillus subtilis σα- dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. Nucleic Acids Res., 23, 2351–2360.
38. Ozoline,O.N., Deev,A.A. and Arkhipova,M.V. (1997) Non-canonical sequence elements in the promoter structure. Cluster analysis of promoters recognized by Escherichia coli RNA polymerase. Nucleic Acids Res., 25, 4703–4709.
39. Huerta,M.A. and Collado-Vides.J. (2003) Sigma70 promoters in Escherichia coli: specific transcription in dense regions of overlapping promoter-like signals. J. Mol. Biol., 333, 261–278.
40. Vogel,J., Axmann,I.M., Herzel,H. and Hess,W.R. (2003) Experimental and computer analysis of transcriptional start sites in the cyanobacterium Prochlorococcus MED4. Nucleic Acids Res., 31, 2890–2899.
41. Burt,T., Mitchell,J., Kolb,A., Minchin,S. and Busby,S. (2000) DNA sequence elements located immediately upstream of the –10 hexamer in Escherichia coli promoters: a systematic study. Nucleic Acids Res., 28, 1864–1870.
42. Mitchell,E.R., Cheng,D., Busby,S.J. and Minchin,S.D. (2003) Identification and analysis of ‘extended –10’ promoters in Escherichia coli. Nucleic Acids Res., 31, 4689–4695.
43. Liss,S. and Margašály,H. (1994) Determination of common structural features in Escherichia coli promoters by computer analysis. Eur. J. Biochem., 223, 823–830.
44. Morrison,D.A. and Jaurin,B. (1990) Identification and analysis of ‘extended –10’ promoters in Escherichia coli. Nucleic Acids Res., 31, 4689–4695.
45. Metcalf,W.W. and Wanner,B.L. (1993) Construction of new broad-host range cloning vector pBBR1MCS, carrying different antibiotic resistance cassettes. Gene. 166, 175–176.
46. Keen,N.T., Tamaki,S., Kobayashi,D. and Trollinger,D. (1988) Improved broad-host range plasmids for DNA cloning in gram-negative bacteria. Gene. 70, 191–197.