# The *Escherichia coli cysG* promoter belongs to the 'extended −10' class of bacterial promoters

Tamara BELYAEVA, Lesley GRIFFITHS, Stephen MINCHIN, Jeffrey COLE and Stephen BUSBY*
School of Biochemistry, University of Birmingham, P.O. Box 363, Birmingham B15 2TT, West Midlands, U.K.

The *Escherichia coli cysG* promoter has been subcloned and shown to function constitutively in a range of different growth conditions. Point mutations identify the −10 hexamer and an important 5′-TGN-3′ motif immediately upstream. The effects of different deletions suggest that specific sequences in the −35 region are not essential for the activity of this promoter *in vivo*. This conclusion was confirmed by *in vitro* run-off transcription assays. The DNAase I footprint of RNA polymerase at the *cysG* promoter reveals extended protection upstream of the transcript start, and studies with potassium permanganate as a probe suggest that the upstream region is distorted in open complexes. Taken together, the results show that the *cysG* promoter belongs to the 'extended −10' class of promoters, and the base sequence is similar to that of the *P1* promoter of the *E. coli* galactose operon, another promoter in this class. *In vivo*, messenger initiated at the *cysG* promoter appears to be processed by cleavage at a site 41 bases downstream from the transcript start point.

## INTRODUCTION

The activity of most *Escherichia coli* promoters is dependent on two hexamer sequences located 10 and 35 bp upstream from the transcription start point [usually referred to as the −10 and −35 sequences; reviewed by McClure (1985)]. However, a small number of promoters have now been identified where specific sequences in the −35 region are not important (Ponnambalam et al., 1986, 1988; Keilty and Rosenberg, 1987; Kumar et al., 1993). At these promoters, RNA polymerase recognizes an extended −10 sequence, and, in all such cases, the motif 5′-TG-3′ is found 1 bp upstream of the −10 hexamer. On the basis of biochemical studies with one particular 'extended −10' promoter, *galP1*, we suggested that the overall architecture of open complexes with RNA polymerase at such promoters might differ from that found with 'consensus promoters', dependent on both −10 and −35 sequences (Busby et al., 1987; Chan and Busby, 1989; Chan et al., 1990; Minchin and Busby, 1993). At *galP1*, RNA polymerase makes little or no contact with the −35 region, but does make extended upstream contacts. Evidence that the upstream region is distorted led us, and others, to suggest that the distortion provides an alternative mechanism for promoter activation in the absence of the normal −35 region contact (Grimes et al., 1991; Lavigne et al., 1992). However, to date, *galP1* is the only extended −10 promoter for which a full genetic and biochemical characterization has been made, and it is not known whether open complexes at other members of this class adopt an unusual architecture with extended upstream contacts. Additionally, it is not clear whether extended −10 promoters are widespread, although the computer-assisted search performed by Kumar et al. (1993) suggested that they are few and far between, at least in the *E. coli* genome.

In this paper we present a genetic and biochemical investigation of the *E. coli cysG* promoter (*pcysG*) and we show that, by several criteria, it is an extended −10 promoter. We discovered this promoter during a study of the *E. coli nirB–cysG* operon, encoding NADH-dependent nitrate reductase (EC 1.6.6.4) which contains two subunits encoded by *nirB* and *nirD* and a sirohaem cofactor (Peakman et al., 1990a,b, Harborne et al., 1992). We located an internal promoter, *pcysG*, just upstream of the *cysG* gene, which encodes a multifunctional enzyme responsible for the biosynthesis of sirohaem from uroporphoryrinogen III (Warren et al., 1990, 1993). Here we show that replacement of the *pcysG* −35 region does not inactivate the promoter, but we have found promoter 'down' mutations in the −10 hexamer and in the 5′-TG-3′ motif located just upstream. *In vitro* studies of open complexes at *pcysG* demonstrate that RNA polymerase makes extended contacts with the upstream sequences and provide evidence for distortion of this region.

## EXPERIMENTAL

### Strains and plasmids

Table 1 lists the *E. coli* strains and bacterial plasmids used in this work, and also the *Eco*RI–*Hin*dIII fragments carrying the *cysG* promoter described here. Standard methods for cloning and manipulating recombinant DNA were used (Maniatis et al., 1982). The starting point for this work was a clone carrying the complete *nirB–cysG* operon (Peakman et al., 1990a) and the plasmid pPMB80 (Harborne et al., 1992). In previous publications the operon sequence was numbered from 1 to 5613, starting at the *Bst*EII site upstream of *nirB* (Figure 1a). However, in the present work we have numbered sequences starting at the *cysG* transcription start point, employing the usual convention of labelling upstream sequences with a '−' prefix and downstream sequences as '+'.

### Manipulation of promoter DNA

The *cys* Δ135, Δ051, Δ022, Δ009 and D+19 fragments were derived from the PMB80 fragment after *Bal*31 exonuclease digestion from the upstream *Eco*RI site and placing the *Eco*RI linker sequence (5′-GAATTCCC-3′) upstream of the truncated *pcysG* sequence, using the protocol described by Busby et al. (1983). The p12C and p14C mutations were created in the *cysG* Δ135 fragment using synthetic oligodeoxynucleotides and the Amersham site-directed mutagenesis kit (cat. no. RPN 1523). To monitor the *in vivo* promoter activities of the *pcysG* derivatives,

* To whom correspondence should be addressed.

**Table 1 Strains, plasmids and promoter fragments used in this work**

| | Details |
|---|---|
| Strains of *E. coli* | |
| M182 | K12 Δ*lac* strain (Casadaban and Cohen, 1980) |
| pop2094 | K12 host facilitating transfer of promoter fragments to bacterial chromosome and fusion of promoter to *lac* (Gaston et al., 1988) |
| Plasmids | |
| pBR322 | Vector for cloning *Eco*RI–*Hin*dIII fragments |
| pOM44 | Vector for cloning *Eco*RI–*Hin*dIII fragments, facilitating transfer of fragments upstream of *lac* genes in pop2094 (Gaston et al., 1988) |
| pAA2053 | *lac* fusion vector (Harborne et al., 1992) |
| pPMB80 | *Eco*RI–*Hin*dIII fragment carrying *pcysG* sequence from co-ordinate 3728 to 4395 (see Figure 1a) cloned into pAA2053 giving a *cysG–lacZ* protein fusion (Harborne et al., 1992) |
| Promoter fragments (all on *Eco*RI–*Hin*dIII fragments) | |
| *pcysG* PMB80 | Fragment carrying *pcysG* sequence from co-ordinate 3728 to 4395 (−421 to +248 with respect to transcript start point) with *Eco*RI linker upstream of −421 and *Hin*dIII linker downstream of +248 (Harborne et al., 1992) |
| *pcysG* Δ135 | Derivative of PMB80 fragment carrying *pcysG* sequence from −135 to +248 with *Eco*RI linker upstream of −135 and *Hin*dIII linker downstream of +248 (this work; Figure 1b) |
| *pcysG* Δ051 | Derivative of Δ135 fragment with *Eco*RI linker upstream of −51 and *Hin*dIII linker downstream of +248 (this work) |
| *pcysG* Δ022 | Derivative of Δ135 fragment with *Eco*RI linker upstream of −22 and *Hin*dIII linker downstream of +248 (this work) |
| *pcysG* Δ009 | Derivative of Δ135 fragment with *Eco*RI linker upstream of −09 and *Hin*dIII linker downstream of +248 (this work) |
| *pcysG* Δ+19 | Derivative of Δ135 fragment with *Eco*RI linker upstream of +19 and *Hin*dIII linker downstream of +248 (this work) |
| *pcysG* Δ135 *p12C* | Derivative of Δ135 fragment with *TA* to *CG* mutation at −12 (this work; Figure 1b) |
| *pcysG* Δ135 *p14C* | Derivative of Δ135 fragment with *GC* to *CG* mutation at −14 (this work; Figure 1b) |

the different fragments were cloned into pAA2053 to give *cysG–lacZ* protein fusions under the control of *pcysG*. β-Galactosidase activities were measured as previously (Harborne et al., 1992) in M182 host cells grown in L broth containing 80 μg/ml ampicillin. The activity of *pcysG* in response to different growth conditions was measured after transfer of the *cysG* Δ135 fragment to a location immediately upstream of a *malQ–lacZ* fusion on the chromosome of strain pop2094, using the protocols detailed by Gaston et al. (1988). β-Galactosidase activities were used to monitor *pcysG* activity as described in Table 2.

### *In vitro* studies

Different *Eco*RI–*Hin*dIII fragments carrying *pcysG* were cloned into pBR322 and preparations were made from caesium chloride gradients. For these studies, complexes between holo-RNA polymerase and DNA fragments were formed in standard buffer containing 20 mM Tris/HCl, pH 8.0, 100 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 1 mM dithiothreitol, 50 μg/ml BSA and 5% glycerol. Run-off transcription assays were performed on purified fragments using the basic protocol of Spassky et al. (1984), described in full detail by Busby et al. (1993). Briefly, 100 nM holo-RNA polymerase was incubated with 10 nM of the fragment and, after open complex formation, transcription was initiated by the addition of $^{32}$P-labelled nucleoside triphosphates plus heparin. Polymerase runs from the open complex to the end of the fragment, giving a discrete band that is detected by autoradiography of a calibrated standard sequence gel. In the kinetic experiment in Figure 2(b), RNA polymerase was incubated with fragments for different periods of time prior to addition of the nucleoside triphosphate/heparin mix. The amount of run-off transcript indicates the extent of open complex formation, since the heparin prevents more than one round of transcription per template.

For footprinting, we adopted the primer extension strategy described by Sasse-Dwight and Gralla (1991). Purified pBR322 DNA, carrying the different *Eco*RI–*Hin*dIII *cysG* promoter fragments, was linearized with *Pst*I. After incubation of 10 nM fragments with 80 nM holo-RNA polymerase, complexes were treated with DNAase I or potassium permanganate as described previously (Chan et al., 1990). After phenol extraction, DNA samples were annealed to the pBR322 'clockwise' *Eco*RI primer, 5′-AGGCGTATCACGAGGCCCT-3′, which had been 5′-end-labelled with polynucleotide kinase and [γ-$^{32}$P]ATP. This primer anneals to sequence just upstream of the *Eco*RI site. Extension with DNA polymerase Klenow fragment proceeds into the *pcysG* promoter but is terminated at positions where the template strand has been modified or cleaved. Sites of termination were deduced after autoradiography of a calibrated sequence gel. For calibration we used a sequence ladder with the same labelled *Eco*RI clockwise primer and unmodified DNA template.

### Transcript mapping

The S1 nuclease method was used to locate the 5′ end of messengers exactly as before (Peakman et al., 1990b). As a probe we used the *pcysG* PMB80 *Eco*RI–*Hin*dIII fragment, 5′-end-labelled with [γ-$^{32}$P]ATP at the *Hin*dIII site. Fragments protected from S1 digestion by RNA were sized on a calibrated sequence gel.

## RESULTS AND DISCUSSION

### Sub-cloning and activity of the *E. coli cysG* promoter

We previously reported the cloning, sequence determination and characterization of the *E. coli nirB–cysG* operon, illustrated in Figure 1(a). We showed that the order of genes in this operon is *nirB-nirD-nirE-nirC-cysG* and that transcription of the entire operon is induced during anaerobic growth due to an FNR-dependent promoter, *pnirB*, located just upstream of the *nirB* open reading frame (Macdonald and Cole, 1985; Peakman et al., 1990a,b; Harborne et al., 1992). However, Peakman et al. (1990b) demonstrated the existence of a second promoter initiating transcription at co-ordinate 4149, in the *nirC* open reading frame and reading into *cysG*. Harborne et al. (1992) showed that this promoter (*pcysG*) was carried on the *Bst*EII–*Hin*cII fragment from co-ordinates 3728 to 4395, and suggested that this promoter is responsible for maintaining *cysG* expression in a wide variety of growth conditions, including aerobic conditions when *pnirB* is inactive.

The starting point of this work was the plasmid pPMB80 (Harborne et al., 1992), which contains the *Bst*EII–*Hin*cII fragment cloned with an *Eco*RI linker upstream of the *Bst*EII site and a *Hin*dIII linker downstream of the *Hin*cII site. The PMB80 fragment carries *nirB–cysG* operon sequence from −421 to +248 with respect to the *pcysG* transcript start point (in this paper we have numbered sequences with respect to the *pcysG* transcript start). Using *Bal*31 nuclease we deleted increasing lengths of sequence from upstream of *pcysG*, and constructed five derivatives of the PMB80 fragment, Δ135, Δ051, Δ022, Δ009 and Δ+19, in which the *Eco*RI linker was placed at positions −135, −51, −22, −9 and +19 respectively (Table 1).

In preliminary studies we found that the Δ135 fragment was fully active and we have used this fragment (Figure 1b) to study

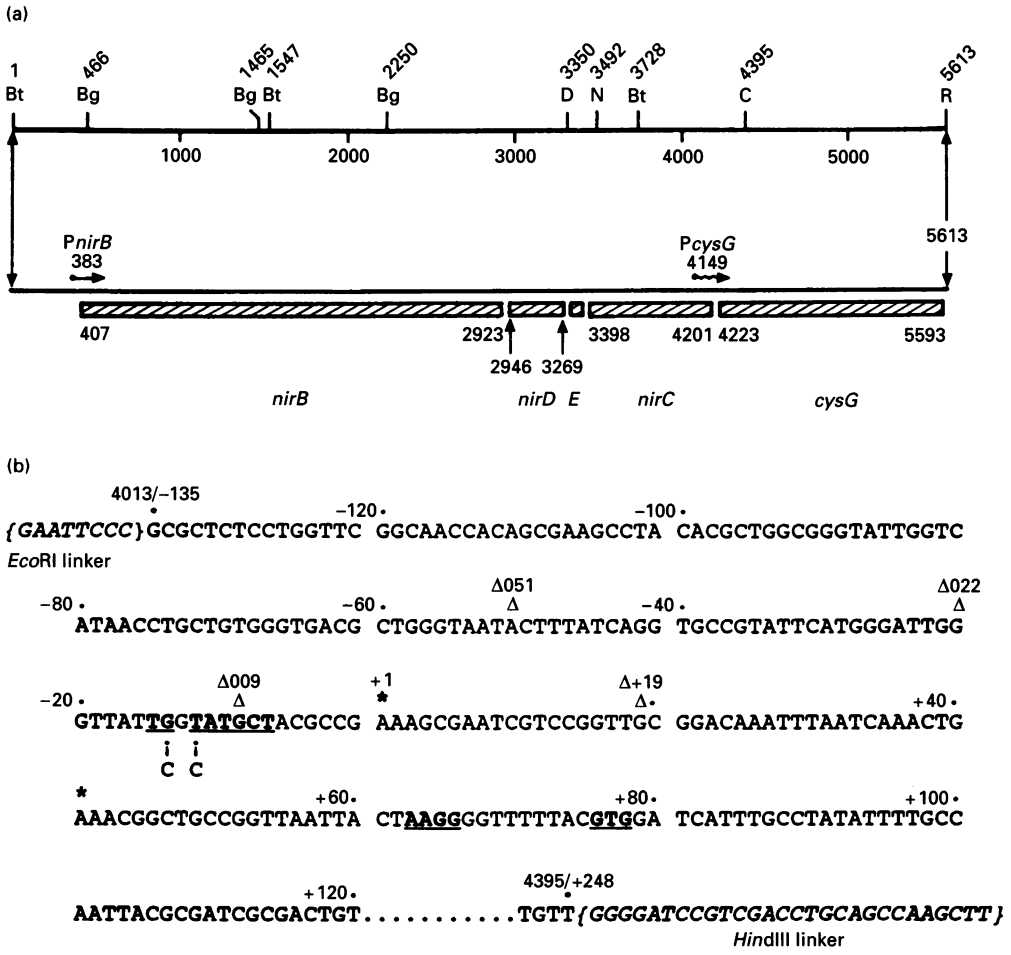


**Figure 1 DNA used in this study**

(**a**) The top line shows a diagram of the 5613 bp *Bst*EII–*Eco*RI fragment carrying the *E. coli nirB–cysG* region numbered as in Peakman et al. (1990a), with the upstream *Bst*EII site taken as co-ordinate 0. The locations of the *nirB, nirD, nirE, nirC* and *cysG* open reading frames are shown as hatched boxes and the translation start and stop sites are numbered. The horizontal arrows depict the *nirB* and *cysG* promoters initiating transcription at positions 383 and 4149. The locations and co-ordinates of a number of restriction sites are shown: Bt, *Bst*EII; Bg, *Bgl*II; D, *Dra*I; N, *Nco*I; C, *Hin*cII; R, *Eco*RI. (**b**) Nucleotide sequence of the upper strand of the Δ135 *Eco*RI–*Hin*dIII fragment carrying sequences from −135 to +248 with respect to the *cysG* transcript start (co-ordinates 4013–4395 of the *nirB–cysG* operon). The asterisk at +1 denotes the *cysG* promoter transcript start site and the second asterisk at +41 shows the 5′ end of the shorter messenger species found *in vivo*. The *pcysG* −10 hexamer and upstream TG motif is in bold type and underlined. The *cysG* Shine–Dalgarno sequence, AAGG, and the translation start codon, GTG, are also emboldened/underlined. The location of the *p12C* and *p14C* mutations at −12 and −14 are shown and the limits of upstream deletions in the Δ051, Δ022, Δ009 and Δ+19 fragments are indicated. The upstream and downstream linker sequences are shown in italics.

the *cysG* promoter *in vivo* and *in vitro*. To investigate possible regulation, the Δ135 fragment was cloned into pOM44 and then transferred to the chromosome of strain pop2094 upstream of a *malQ–lacZ* fusion. Expression of this fusion under the control of *pcysG* was then measured during aerobic or anaerobic growth in a variety of media (Table 2). We previously found that *pcysG* activity was insensitive to nitrite or nitrate availability (Peakman et al., 1990b). However, since sirohaem is also essential for sulphite assimilation, we anticipated that the nature of sulphur nutrients might trigger activity changes. However, the data in Table 2 show that *pcysG* activity is unaffected by a range of sulphur-rich and sulphur-poor regimes, and suggest that *pcysG* is a constitutively expressed unregulated promoter.

## Deletions and mutations in the *cysG* promoter

Figure 1(b) shows the nucleotide sequence of the top strand at the *cysG* promoter region in the Δ135 fragment. In previous work, Peakman et al. (1990b) identified a likely −10 hexamer, 5′-TATGCT-3′, but were unable to find a −35 region hexamer resembling the consensus. The 5′-TGN-3′ motif immediately upstream of the −10 hexamer suggested that *pcysG* might be an extended −10 promoter. To investigate this, we created nested deletions from the *Eco*RI linker at −135 that moved the linker to different locations in the promoter. *Eco*RI–*Hin*dIII fragments carrying the truncated *cysG* promoters were cloned into the *lac* expression vector pAA2053 and the resulting recombinants were transformed into the Δ*lac* strain M182. The β-galactosidase activity in transformants carrying the different plasmids was measured (Table 3). Replacement of the −35 region by the deletion to −22 caused no reduction in *pcysG* activity, whilst longer deletions to −9 or +19 stopped expression. To confirm that the −10 hexamer and TGN motif were indeed responsible for *pcysG* activity, we used site-directed mutagenesis to make two single point mutations (*p12C* and *p14C*; see Table 1 and Figure 1b) in the Δ135 fragment. Both of these mutations cause large

**Table 2 β-Galactosidase activity of pop2094 cells carrying a chromosomally encoded *pcysG–lac* fusion**

β-Galactosidase activities were determined using standard protocols (Miller, 1972) and are expressed in Miller units. The strain pop2094 contains a *malQ–lacZ* fusion under the control of the *pcysG* Δ135 fragment. This fragment was cloned into pOM44 and transferred to the pop2094 chromosome as described by Gaston et al. (1988). Cells were grown in the sulphur-free minimal medium described by Cole et al. (1974) with 0.4% glucose as a carbon source supplemented with different sources of sulphur as listed. Since strain pop2094 is *aroB*, the medium was also supplemented with 25 μg/ml tryptophan, phenylalanine, tyrosine and shikimic acid.

| Supplement to media | β-Galactosidase activity (units) | |
|---|---|---|
| | Aerobic growth | Anaerobic growth |
| 10 mM $MgSO_4$ | 220 | 140 |
| 1 mM Djenkolic acid | 130 | 200 |
| 10 mM $MgSO_4$ + 25 μg/ml cysteine | 250 | 120 |
| 10 mM $MgSO_4$ + 25 μg/ml cysteine + 1 mM djenkolic acid | 220 | 200 |

**Table 3 *In vivo* activity of different *pcysG* derivatives**

β-Galactosidase activities; expressed in Miller (1972) units, are the averages of four independent determinations using standard protocols. *Eco*RI–*Hin*dIII fragments carrying different *pcysG* sequences, as listed in Table 1, were cloned into the *lac* fusion vector pAA2053. Activities are taken as a measure of promoter strength, as in our previous work (Harborne et al., 1992). The recombinants were transformed into the Δ*lac* strain M182, and expression of the *cysG–lacZ* fusion protein was measured during growth in L broth containing 80 μg/ml ampicillin. The data refer to aerobic growth; however, expression differed by less than 10% during anaerobic growth.

| *pcysG* promoter fragment fused to *lac* in pAA2053 | β-Galactosidase activity (Miller units) |
|---|---|
| No promoter | 5 |
| Δ135 | 500 |
| Δ051 | 350 |
| Δ022 | 540 |
| Δ009 | 5 |
| Δ+19 | 5 |
| Δ135 *p12C* | 25 |
| Δ135 *p14C* | 60 |

reductions in *pcysG* activity (Table 3), and we conclude that *pcysG* activity depends on an extended −10 promoter but not on specific −35 region sequences.

## *In vitro* transcription from the *cysG* promoter

Figure 2(a) shows the results of run-off transcription assays using purified fragments carrying different *pcysG* derivatives. With the Δ135 fragment, a clear 271 base transcript is seen which corresponds to transcription initiation at +1. This transcript is absent with fragments carrying either the *p12C* or *p14C* mutations, in agreement with results from *in vivo* assays. To demonstrate that the *pcysG* −35 region sequence is not needed, run-off assays were also performed to compare the Δ135 and Δ022 promoters. Clear initiation at +1 was found with both promoters

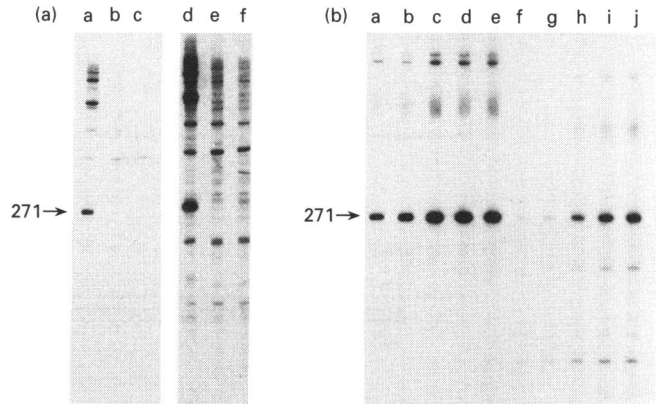


**Figure 2 Run-off transcription assays at the *cysG* promoter**

Autoradiogram of a sequence gel to analyse RNA products made *in vitro*. The 271 nucleotide run-off transcript initiating at +1 in *pcysG* is indicated. The longer bands are due to the run-off transcript running back through the template as described by Spassky et al. (1984). (**a**) Run-off transcription with the *Eco*RI–*Hin*dIII Δ135 fragment carrying the wild-type *pcysG* sequence (lanes a and d), the *p12C* (lanes b and e) or the *p14C* (lanes c and f) mutations. Lanes d–f show an overexposure of the gel in lanes a–c. (**b**) Run off transcription assays with the *Eco*RI–*Hin*dIII Δ135 fragment carrying the wild-type *pcysG* sequence (lanes a–e) or the *Hae*III–*Hin*dIII fragment carrying the *cysG* promoter truncated with the Δ022 deletion (*Hae*III cuts just upstream of the *Eco*RI site in pBR322). To measure the kinetics of open complex formation, DNA fragments were incubated with RNA polymerase for different times prior to initiating transcript extension: lanes a and f, 1 min; lanes b and g, 2 min; lanes c and h, 4 min; lanes d and i, 8 min; lanes e and j, 16 min.

(Figure 2b). However, in our experimental conditions, open complex formation at the truncated promoter is four times slower than with the longer fragment. This demonstrates that the upstream sequences do play a role, although they are not essential for promoter activity, as was found in a similar study with *galP1* (Chan and Busby, 1989). In this experiment we consistently failed to obtain full occupancy of *pcysG* with the truncated promoter, presumably because the weakened promoter is less competitive against non-productive non-specific binding sites that sequester the RNA polymerase.

## Footprint analysis at the *cysG* promoter

Open complexes between RNA polymerase and the *cysG* promoter were made using the Δ135 fragment cloned in pBR322 and footprint analysis was performed using DNAase I. RNA polymerase gave a clear extended footprint from +20 to around −60 (Figure 3, lanes a–c), similar to that seen at *galP1*. In a control experiment with the same fragment carrying the *p12C* mutation, no footprint was found (Figure 3, lanes d–f).

The complexes were also probed with potassium permanganate to monitor duplex distortions. The results in Figure 4 show that RNA polymerase induces reactivity of a number of bases around −5, due to unwinding close to the transcript start site. A strong polymerase-induced band is also apparent around −50. Chan et al. (1990) found a similar band at *galP1* and interpreted this as polymerase-induced distortion of upstream sequences, although it cannot be proved formally that it is not due to unwinding. Unwinding about the *pcysG* transcript start is suppressed by the *p12C* mutation but the polymerase-induced enhancement around −50 is increased (Figure 4, lanes d–f). To explain this, we suggest that contacts between RNA polymerase and upstream sequences are possible even in the absence of open complex formation;

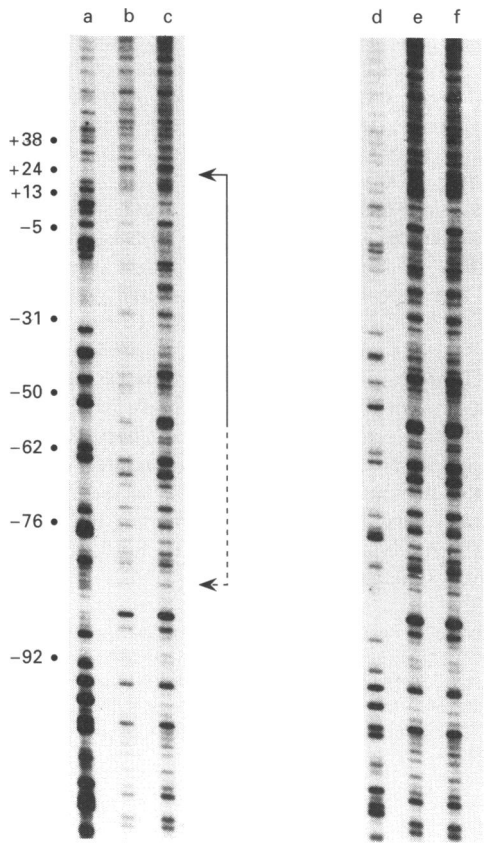


**Figure 3 Footprint analysis of open complexes at the *cysG* promoter**

Autoradiograms of gels run to analyse primer extension products after DNAase I digestion of complexes between RNA polymerase and the Δ135 fragment carrying wild-type *pcysG* (lanes a–c) or *pcysG p12C* (lanes d–f). Digestion was performed on DNA incubated either with (lanes b and e) or without (lanes c and f) RNA polymerase. Sequencing runs were loaded in lanes a and d to give the calibration. The arrows show the region protected from digestion by RNA polymerase.

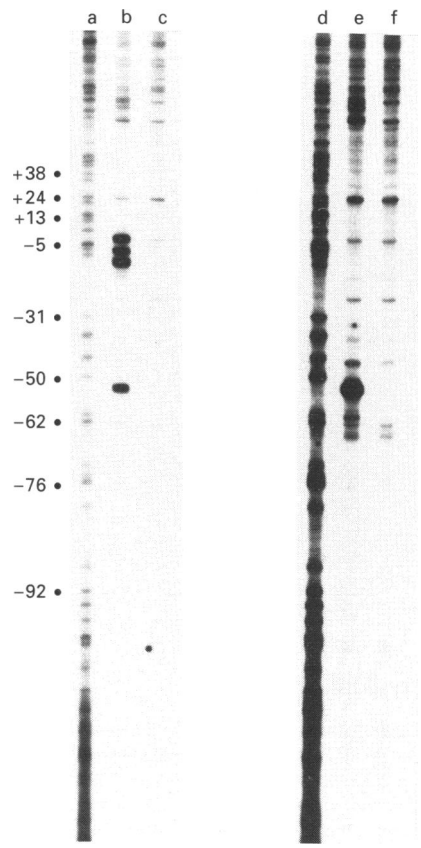


**Figure 4 Unwinding and distortion in open complexes at the *cysG* promoter**

Autoradiograms of gels run to analyse primer extension products after potassium permanganate treatment of complexes between RNA polymerase and the Δ135 fragment carrying wild-type *pcysG* (lanes a–c) or *pcysG p12C* (lanes d–f). Digestion was performed on DNA incubated either with (lanes b and e) or without (lanes c and f) RNA polymerase. Sequencing reactions were loaded in lanes a and d to give the calibration.

similar contacts have been noted with some *galP1* derivatives (Johnston et al., 1987; Chan et al., 1990).

### *In vivo* analysis of the *cysG* transcript start

Although our present results suggest that *pcysG* is a 'simple' extended −10 promoter, our previous work, using S1 nuclease to map the transcript start *in vivo*, revealed two 5′-ends, one located at +1 and the other at +41 (Figure 1b; Peakman et al., 1990b). To explain this, we suggested that the 5′-end at +41 is generated by processing of the longer message starting at +1, although we could not rule out the possibility of a second *cysG* promoter. The experiments reported in the present paper argue for processing and against a second promoter. Firstly, there is no *in vitro* evidence for initiation at +41 in either the run-off or footprinting experiments. Secondly, the *p12C*, *p14C* and Δ009 deletions suppress all *pcysG* activity, suggesting that +1 is the sole transcript initiation site. To corroborate this, we repeated the S1 transcript mapping experiment to investigate the effects of different deletions. Total RNA was extracted from cells containing pAA2053 carrying the PMB80, Δ051, Δ022, Δ009 or Δ+19 *pcysG* fragments. The 5′-ends of *pcysG* mRNAs were mapped by the S1 nuclease method using the end-labelled *Eco*RI–*Hin*dIII PMB80 fragment as a probe. With RNA from cells carrying the PMB80 fragments, protected bands corresponding to 5′-ends at +1 and +41 ($S_{+1}$ and $S_{+41}$) are seen (Figure 5, lane b), in agreement with Peakman et al. (1990b). Note that some full-length protected probe is seen at the top of lane b, which is due to a promoter in the vector reading through across the *Eco*RI site into the PMB80 fragment. With RNA from cells carrying the Δ051 or Δ022 fragments, bands $S_{+1}$ and $S_{+41}$ are still seen, but the read-through bands W and X appear to be shorter (Figure 5, lanes c and d). With RNA from cells carrying the Δ009 or Δ+19 fragments, the read-through bands, Y and Z, are again shorter, but there is a substantial parallel reduction in the intensity of the $S_{+1}$ and $S_{+41}$ bands (lanes e and f). The simplest explanation for these results is that the $S_{+41}$ transcript is derived from the $S_{+1}$ transcript by processing. Note, however, that a faint $S_{+41}$ band is seen in both lanes e and f, and so the possibility of a second downstream weak promoter cannot be excluded (as is the situation with *gal*; Ponnambalam et al., 1988). Alternatively, these faint bands could be due to processing of the longer read-through transcript.

### Conclusions

Although the function of the NADH-dependent nitrite reductase is the dissimilation of nitrite during anaerobic growth of *E. coli*

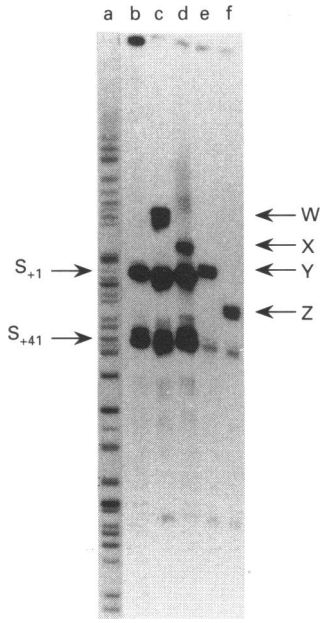


**Figure 5 Analysis of *cysG* transcripts *in vivo***

Sequence gel to analyse probe fragments protected from S1 nuclease digestion by RNA extracted from cells containing pAA2053 carrying different *pcysG* promoter fragments: lane b, PMB80; lane c, Δ051; lane d, Δ022; lane e, Δ009; lane f, Δ+19. The probe for each lane was the *Eco*RI–*Hin*dIII PMB80 fragment end-labelled at the *Hin*dIII site, and the gel was calibrated with a Maxam–Gilbert C reaction on the probe. $S_{+1}$ and $S_{+41}$ identify protected fragments due to RNA species with 5′-ends at +1 and +41 respectively. W, X, Y and Z identify bands due to transcripts originating in the plasmid vector.

(Cole, 1988), the *cysG* product is required during both aerobic and anaerobic growth, since sirohaem is an essential cofactor during the reduction of both nitrite and sulphite. Hence, although *cysG* can be expressed from the *nirB* promoter, it can also be transcribed from *pcysG*, which appears to function constitutively. Unusually, *pcysG* sequences fall in a functional open reading frame, *nirC*, and presumably this places constraints on possible promoter sequences. It is tempting to suppose that this is why an extended −10 promoter is used here, since this type of promoter contains less sequence information than those dependent on both −10 and −35 hexamers. Figure 6(a) shows a comparison of the *E. coli pcysG* sequence with the corresponding region of the closely related *Salmonella typhimurium* chromosome, in which the *nirC* and *cysG* genes have been sequenced recently (Wu et al., 1991). The sequences upstream of +1 are well conserved. In contrast, there is greater divergence in base sequences around +40, suggesting that any weak promoters in this downstream region have little, if any, importance.

The results presented here strongly argue that the *cysG* promoter is a member of the extended −10 class and does not require specific −35 region sequences. To date, the only other promoter of this type that has been thoroughly characterized both genetically and biochemically is *galP1*. Our results suggest a number of features common to both promoters, including an important TG motif upstream of both −10 hexamers, extended upstream contacts with RNA polymerase in the open complexes, and evidence for distortion of upstream sequences. Remarkably, the *galP1* and *cysG* promoter sequences are related: Figure 6(b) shows that the promoters share common bases at 24 of the positions from +1 to −50. It will be interesting to compare the kinetic and thermodynamic parameters of both promoters to identify further similarities.

This work was funded by grants from the U.K. Science and Engineering Research Council, the Royal Society and the Society for General Microbiology. T.B. was supported by a Royal Society grant for scientists from Russia and S.B. is an EPA Cephalosporin Fund Research Fellow. We thank Helen Burns for help in developing the primer extension technology, and Richard Hayward for communicating his results prior to publication.

## REFERENCES

Busby, S., Kotlarz, D. and Buc, H. (1983) J. Mol. Biol. **167**, 259–274
Busby, S., Spassky, A. and Chan, B. (1987) Gene **53**, 145–152
Busby, S., Kolb, A. and Minchin, S. (1993) in Protocols in Molecular Biology (Kneale, G. G., ed.), chapter 31, Humana Press, Totawa, NJ
Casadaban, M. and Cohen, S. (1980) J. Mol. Biol. **138**, 179–207
Chan, B. and Busby, S. (1989) Gene **84**, 227–236
Chan, B., Spassky, A. and Busby, S. (1990) Biochem. J. **270**, 141–148
Cole, J. (1988) Symp. Soc. Gen. Microbiol. **42**, 281–329
Cole, J., Coleman, K., Compton, B., Kavanagh, B. and Keevil, C. (1974) J. Gen. Microbiol. **85**, 11–22
Gaston, K., Chan, B., Kolb, A., Fox, J. and Busby, S. (1988) Biochem. J. **253**, 809–818
Grimes, E., Busby, S. and Minchin, S. (1991) Nucleic Acids Res. **19**, 6113–6118
Harborne, N., Griffiths, L., Busby, S. and Cole, J. (1992) Mol. Microbiol. **6**, 2805–2813
Johnston, F., Ponnambalam, S. and Busby, S. (1987) J. Mol. Biol. **195**, 745–748
Keilty, S. and Rosenberg, M. (1987) J. Biol. Chem. **262**, 6389–6395
Kumar, A., Malloch, R., Fujita, N., Smillie, D., Ishihama, R. and Hayward, R. (1993) J. Mol. Biol. **232**, 406–418

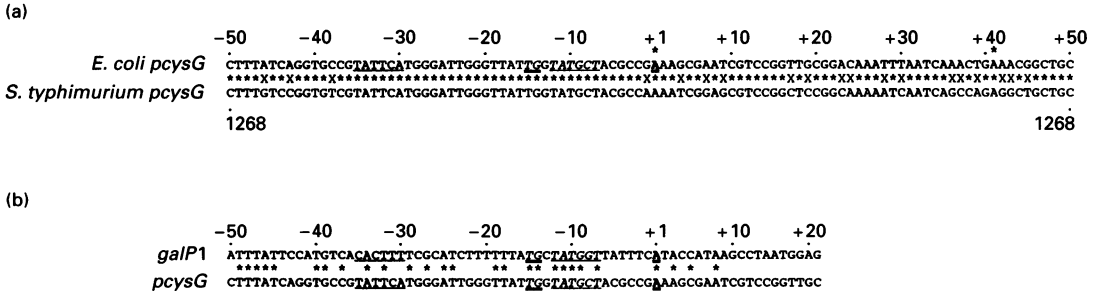


**Figure 6 Sequence comparisons**

(**a**) The nucleotide sequence of the top strand of the *E. coli cysG* promoter from −50 to +50 is compared with the cognate region from *S. typhimurium* (Wu et al., 1991). Asterisks indicate identical bases and X denotes a difference. (**b**) The nucleotide sequences of the top strands of the *E. coli cysG* and *galP1* promoters from −50 to +20 are compared, aligned by the transcript start at +1 and the −10 hexamers. Asterisks denote identical bases. Note that the *galP1* sequence shown here carries the *p197* mutation that inactivates the alternative *galP2* promoter (Busby et al., 1987).

Lavigne, M., Kolb, A. and Buc, H. (1992) Biochemistry **31**, 9647–9656
Macdonald, H. and Cole, J. (1985) Mol. Gen. Genet. **200**, 328–334
Maniatis, T., Fritsch, E. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
McClure, W. (1985) Annu. Rev. Biochem. **54**, 171–204
Miller, J. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Minchin, S. and Busby, S. (1993) Biochem. J. **289**, 771–775
Peakman, T., Crouzet, J., Mayaux, J.-F., Busby, S., Mohan, S., Harborne, N., Wootton, J., Nicolson, R. and Cole, J. (1990a) Eur. J. Biochem. **191**, 315–323

Received 5 July 1993/20 August 1993; accepted 24 August 1993

Peakman, T., Busby, S. and Cole, J. (1990b) Eur. J. Biochem. **191**, 325–331
Ponnambalam, S., Webster, C., Bingham, A. and Busby, S. (1986) J. Biol. Chem. **261**, 16043–16048
Ponnambalam, S., Chan, B. and Busby, S. (1988) Mol. Microbiol. **2**, 165–172
Sasse-Dwight, S. and Gralla, J. (1991) Methods Enzymol. **208**, 146–168
Spassky, A., Busby, S. and Buc, H. (1984) EMBO J. **3**, 43–50
Warren, M., Stolowich, N., Santander, P., Roessner, C., Sowa, B. and Scott, A. (1990) FEBS Lett. **261**, 76–80
Warren, M., Bolt, E. and Woodcock, S. (1993) CIBA Symp. **180**, in the press
Wu, J.-Y., Siegel, L. and Kredich, N. (1991) J. Bacteriol. **173**, 325–333