Transcript Analysis of the Citrate Synthase and Succinate Dehydrogenase Genes of Escherichia coli K12

By ROBIN J. WILDE† AND JOHN R. GUEST*
Department of Microbiology, Sheffield University, Sheffield S10 2TN, UK

(Received 19 May 1986; revised 9 July 1986)

A transcript analysis of the citrate synthase and succinate dehydrogenase genes (gltA–sdhCDAB) of Escherichia coli was done by nuclease S1 mapping. Evidence was obtained for two monocistronic gltA transcripts extending anti-clockwise, to a common terminus, from independent promoters with start points 196 bp (major) and 299 bp (minor) upstream of the gltA coding region. Evidence was also obtained for two polycistronic sdh transcripts, sdhCDAB (minor) and sdhDAB (major), extending clockwise, from sites 219 bp upstream of sdhC and 1455 bp upstream of sdhD (i.e. within sdhC), to a common terminus. The synthesis of all of the transcripts was repressed by growth in the presence of glucose, and this is consistent with the well-established fact that both enzymes are subject to catabolite repression. Sequences resembling known binding sites for the cAMP–CRP (cyclicAMP–cyclicAMP receptor protein) complex occur in the vicinity of each promoter suggesting that they are activated by the cAMP–CRP complex.

INTRODUCTION

Citrate synthase (EC 4.1.3.7) catalyses the condensation of acetyl-CoA with oxaloacetate, in the first step of the citric acid cycle. It performs important metabolic functions in energy production and biosynthesis, and its key position has prompted detailed investigations of the structural, kinetic and regulatory properties of the enzymes from diverse sources (Srere, 1975; Weitzman, 1981). The citrate synthase of Escherichia coli typifies those of Gram-negative bacteria in being a 'large' enzyme comprising six identical subunits of 48 kDa. Its activity is inhibited by NADH, ATP and 2-oxoglutarate, products that are thought to regulate the metabolic flux through the cycle.

Succinate dehydrogenase (EC 1.3.99.1) is another citric acid cycle enzyme; it catalyses the interconversion of succinate and fumarate and donates electrons to the aerobic respiratory chain (Hederstedt & Ruthberg, 1981). Recent studies, including nucleotide sequencing plus in vitro translation of the cloned genes and specific immunoprecipitation of the enzyme complex, have revealed four subunits, the flavoprotein (64 kDa), iron-sulphur protein (27 kDa) and two hydrophobic subunits (14 and 13 kDa), encoded by the sdhA, B, C and D genes, respectively (Wood et al., 1984; Darlison & Guest, 1984; Condon et al., 1985). Furthermore, it now appears that one of the hydrophobic subunits is cytochrome b556, i.e. the sdhC and cybA genes are identical (Murakami et al., 1985).

The citrate synthase (gltA) and succinate dehydrogenase (sdhCDAB) genes are located at 16.5 min in the E. coli K12 linkage map, in a cluster encoding two other citric acid cycle enzymes: the 2-oxoglutarate dehydrogenase complex (E1o and E2o components; sucAB) and succinyl-CoA synthetase (sucCD). The successful isolation of the gltA gene (Guest, 1981) led to the cloning and identification of the adjacent genes (Spencer & Guest, 1982; Buck et al., 1986).

† Present address: Leicester Biocentre, University of Leicester, University Road, Leicester LE1 7RH, UK.
and the nucleotide sequence of the entire segment (13061 bp) has now been established (Hull et al., 1983; Ner et al., 1983; Wood et al., 1984; Darlison & Guest, 1984; Darlison et al., 1984; Spencer et al., 1984; Buck et al., 1985). The gene organization and transcriptional polarities are as follows:

\[ \text{gltA-sdhCDAB-sucAB-sucCD} \]

and it would appear that transcription of the \( \text{gltA} \) and \( \text{sdh} \) genes diverges from a common 707 bp intergenic region containing two pairs of potential promoter sequences (Wood et al., 1984).

As the first enzyme in the citric acid cycle, it has been generally assumed that citrate synthase plays a major role in controlling the carbon flow through the cycle. However, recent studies in which the enzyme was expressed at 0.1\text{-} to 50-times the normal level, by fusing the \( \text{gltA} \) gene to a controllable promoter, have shown that this is not always the case (Walsh & Koshland, 1985). When acetate is the sole carbon source, the carbon flow through the cycle and the growth rate are directly related to the citrate synthase level, but with glucose plus acetate the enzyme is normally expressed at levels well in excess of the flux- and growth-limiting amount. Relatively little is known about the mechanisms controlling transcription of the \( \text{gltA} \) gene, although measurements of citrate synthase protein and activity in cultures of \( \text{E. coli} \) grown in different media have shown that it resembles other citric acid cycle enzymes in being repressed by glucose and anaerobiosis, and induced by acetate and oxygen, and its synthesis is inversely related to growth rate (Gray et al., 1966; Smith & Neidhardt, 1983). Succinate dehydrogenase is likewise induced aerobically, but repressed anaerobically and by glucose. In this respect it differs from the analogous membrane-bound flavoprotein complex, fumarate reductase, which is the terminal reductase of a specific anaerobic electron transport chain that is induced anaerobically and repressed by oxygen or nitrate (Spencer & Guest, 1973, 1974; Ruiz-Herrera & Garcia, 1972; Takahashi, 1975; Cole et al., 1985).

In this paper we report a transcript analysis of the \( \text{gltA-sdhCDAB} \) region using nuclease S1 mapping, as a first step in defining the molecular mechanisms controlling the synthesis of citrate synthase and succinate dehydrogenase. Two monocistronic \( \text{gltA} \) transcripts extending from separate promoters to a termination site just downstream of the \( \text{gltA} \) structural gene were identified. Evidence was also obtained for two polycistronic \( \text{sdh} \) transcripts encoding the \( \text{sdhCDAB} \) and \( \text{sdhDAB} \) genes. All four transcripts were subject to catabolite repression.

**METHODS**

*Bacteria and plasmids.* Two strains of \( \text{E. coli} \) K12 were used: ED8641 (\( \text{supE} hsdR \text{recA56} \)), an auxotrophic transformable host for plasmids pGS91 and pGS65 (=pGS141); and JM101 (\( \text{thi supE proA} \text{B-lac} \text{O/F traD proAB*lacZAMIS} \text{lacI} \)), the transfection host used for propagating M13 phages (Sanger et al., 1980). Plasmids pGS91 and pGS65 are derivatives of pBR322 containing 5.07 kb \( \text{BgII-XhoI} \) and 4.5 kb \( \text{BamHI} \) (\( \text{B}_{1}-\text{B}_{2} \)) fragments and conferring \( \text{GltA}^{+} \text{Amp}^{\text{R}} \) (\( \text{Tet}^{\text{S}} \)) and \( \text{Sdh}^{+} \text{Amp}^{\text{R}} \) (\( \text{Tet}^{\text{S}} \)) phenotypes, respectively (Hull et al., 1983; Wood et al., 1984).

**M13 probes.** The probes were constructed by cloning specific segments of \( \text{E. coli} \) DNA in M13 vectors. Probes A and A' were obtained by 'shot-gun' cloning an \( \text{MslI} \) digest of the 1.27 kb \( \text{NruI-HindIII} \) fragment of pGS91 into \( \text{AccI} \)-cut M13mp11, so that the 456 bp \( \text{MslI} \) fragment was recovered in both orientations. Probes B and B' contained the 739 bp \( \text{BamHI-HindIII} \) fragment of \( \text{JG117} \) (Spencer & Guest, 1982) in M13mp10 and M13mp11, respectively. Probes D and D' contained the 1462 bp \( \text{NruI-BglII} \) fragment of \( \text{JG117} \) cloned between the \( \text{Smal} \) and \( \text{BamHI} \) sites of M13mp18 (\( \text{D}^{\text{O}} \)) and M13mp19 (\( \text{D}^{\text{R}} \)). Probes C and C' contained the 660 bp \( \text{Sau3A-BglII} \) fragment inserted in both orientations in the \( \text{BamHI} \) site of M13mp18. This fragment was isolated from a \( \text{Sau3A} \) digest of the double-stranded segment of DNA formed by hybridizing probes D and D' according to Miles & Guest (1985). Probes E, F, G' and H contained inserts of 310, 807, 223 and 351 bp (respectively), derived by cloning ultrasonic fragments of the 3-9 kb \( \text{NruI segment} \) (\( \text{N}^{\text{N}_{0}} \)) of pGS91 into the \( \text{Smal} \) site of M13mp8 (Wood et al., 1984). Two complementary probes (F and G) were made by a 'turn-round' procedure that involved primer extension across the inserts of F and G' according to Hong (1981) and recloning the \( \text{EcoRI-PvuI} \) fragment between the corresponding sites of M13mp19. Probes I and I' were obtained by 'shot-gun' cloning a \( \text{Sau3A} \) digest of the 2.34 kb \( \text{HindIII-Xhol} \) fragment of pGS65 in \( \text{BamHI} \)-cut M13mp8 so that the 634 bp \( \text{Sau3A} \) fragment was recovered in both orientations. Probes spanning the \( \text{sdhB-sucA} \) intergenic region have been described previously by Spencer & Guest (1985): J
and J' contain a 717 bp Sau3A fragment recovered as for I and I' after hybridizing two complementary M13 derivatives (probes K and K') containing the 1.38 kb XhoI–BamHI fragment (X–B.). Single-stranded DNA from the M13 derivatives was prepared according to Sanger et al. (1980) and the presence of the desired inserts was confirmed by sequencing with deoxyadenosine-5'-[35S]thiotriphosphate and salt-gradient gels (Biggin et al., 1983).

Radioactively-labelled M13 probes were prepared by propagating the phage in low-phosphate medium (1-2 ml) containing 10 μCi (370 kBq) [32P]orthophosphate according to Sanger et al. (1982).

**RNA preparation and hybridization.** RNA that is potentially enriched for gltA mRNA and sdh mRNA was prepared by hot phenol extraction at pH 5-0 according to Salser et al. (1967), starting with late exponential phase cultures of ED8641(pGS91) and ED8641(pGSM65). The cultures were grown in L broth (Guest, 1981) supplemented with ampicillin (50 μg ml⁻¹) and glucose (1%, w/v) was added to repress the expression of citrate synthase and succinate dehydrogenase when required.

Nuclease S1 transcript mapping was based on the method of Squires et al. (1981) except that M13 probes were sometimes labelled with 32P. E. coli RNA (200 or 400 μg) and M13 DNA (~ 5 μg) were annealed at 68°C for 30 min and slowly cooled to room temperature before digesting with nuclease S1 (55 units; Vogt, 1980) at 37°C for 4 h. Nuclease-S1-resistant hybrids were precipitated with ethanol and dried under vacuum. The hybrid molecules were analysed by electrophoresis in non-denaturing polyacrylamide (6%, w/v) gels alongside a 35S-labelled sequencing ladder obtained from an appropriate M13 clone using the dyeoxy method (Biggin et al., 1983).

**Screening for conserved sequences.** The gltA–sdh intergenic region was searched for potential promoter sequences resembling the consensus sequence TCTTGACAT—T/ 7 to 9 bp /T–TG–TATAAT (based on the 112 promoters analysed by Hawley & McClure, 1983) using the SEQFIT program of Staden (1977). Each potential promoter was assigned a probability score calculated by multiplying the frequencies given by Hawley & McClure (1983). Both strands of the region were likewise searched for potential CRP-binding sites using the consensus sequence of Chapon & Kolb (1983), AA–TGTGA–TA–TCAC3’/ATT. Special attention was paid to the TGTGA segment which occurs in at least one strand of 19 out of 26 CRP-binding sites (Busby, 1986).

**Materials.** Restriction endonucleases, T₄-DNA ligase, DNA-polymerase (Klenow fragment) and M13 vector replicative form DNA were purchased from Gibco-BRL, Boehringer and Pharmacia-PL. Nuclease S1 was from Boehringer, the 123 bp ladder was from Gibco-BRL, and the radiochemicals, deoxyadenosine 5’-[α-35S] thiotriphosphate (400 Ci mmol⁻¹; 148 TBq mmol⁻¹) and [32P]orthophosphate (carrier free) were obtained from Amersham.

**RESULTS**

**Transcription analysis of the gltA gene**

Transcripts of the gltA region were analysed with RNA prepared from a plasmid-containing strain, ED8641(pGS91), which is potentially enriched for the gltA transcript. Plasmid pGS91 is a gltA* derivative of pBR322 in which a 5-07 kb BglII–XhoI fragment, containing a functional gltA gene and three sdh genes (sdhCDA), is inserted between the BamHI and SalI sites in the tet gene of the vector (Hull et al., 1983; Wood et al., 1984). This plasmid complements the metabolic lesion of gltA recA strains, directs the synthesis of the 48 kDa citrate synthase polypeptide in maxicells, and amplifies citrate synthase activity some 15-fold. The plasmid-encoded gltA gene is expressed from its own promoter in the gltA–sdh intergenic region, with a polarity opposing that of the flanking tet promoter (Fig. 1). The relevant segment of DNA (orientated anticlockwise relative to the linkage map) and the single-stranded M13 probes used in RNA : DNA hybridization are shown in Fig. 1. Four different M13 probes (A, B, C and D) were used to map the gltA transcripts, and four probes (A’, B’, C’ and D’) containing the corresponding complementary strands were used as controls.

**Location of gltA transcription start points.** A typical set of nuclease-S1-resistant RNA : DNA hybrids obtained with different probes are shown in Fig. 2. Probe A generated two hybrids of ~160 bp (major band) and ~263 bp (minor band). These correspond to major and minor transcripts having start points ~196 bp (S₁) and ~299 bp (S₂) upstream of the gltA coding region, respectively (Fig. 1). Probe B produced just one hybrid of 150 bp corresponding to the minor transcript starting ~302 bp upstream of the structural gene. The hybrid formed between
Fig. 1. Diagram showing the single-stranded M13 probes used to map the gltA transcripts. The upper portion shows the coding regions (---) and intergenic regions (— —) in relevant segments of chromosomal and plasmid pGS91 DNA, drawn to scale except across the discontinuities (---). Left to right corresponds to anticlockwise in the E. coli linkage map. Restriction sites are abbreviated as follows: B, BamHI; Bg, BglII; E, EcoRI; H, HindIII; M, MspI; N, NruI; S, Sau3A. The BamHI sites are numbered according to Wood et al. (1984), and the subscript refers to sites in the vector (pBR322). The coordinates for the first base of each restriction site, and the start and stop codons for genes are numbered from the HindIII site according to the nucleotide sequences of Ner et al. (1983) and Wood et al. (1984). The coordinates in parentheses represent those used in describing the sdh transcripts. The transcripts are denoted by wavy lines, with arrows to show their polarity and vertical dashed lines to indicate the start points (S) and termination sites. The lower portion shows the polarities and extents of DNA cloned in the M13 probes. The numbers above the probe lines are the sizes (bp) of the corresponding nuclease S1-resistant RNA:DNA hybrids; the asterisks (*) denote accurate sizes determined in denaturing gels.

The major transcript and probe B is too small (~40 bp) to be stable under the conditions used for digestion. Of the control probes, A' and B', only B' generated a hybrid molecule (~400 bp). This clearly corresponds to the sdh transcript which extends out of the same intergenic region, but in the opposite direction into the sdhCDAB operon (Fig. 1).

Radioactively labelled DNA strands of the S1-resistant hybrids from probe A were sized more accurately by electrophoresis in denaturing gels alongside a sequencing ladder derived from probe A (Fig. 3a). The major band is a 160-mer, which locates the start point (S2) of the major gltA transcript at position 696, exactly 196 bp upstream of the gltA translational initiation codon at position 892 (Fig. 4). The minor band is represented by a cluster of 260 to 263-mers, of which the strongest (the 263-mer) identifies the start point (S1) for the minor transcript as position 593, 299 bp upstream of the structural gene (Fig. 4). The S1-resistant hybrids obtained in transcript mapping often prove to be heterogenous clusters when analysed on high resolution gels. The heterogeneity could be due to the existence of more than one start point or to variability of S1 digestion resulting in the incomplete removal of all unpaired bases or the removal of bases that are actually paired.

Since none of the potential gltA promoters suggested previously by Wood et al. (1984) or Ner et al. (1983) corresponds to the major transcript detected here, the sequence was examined more rigorously using algorithms to quantitate the relationships between potential promoters and established consensus sequences (see Methods). This revealed a promoter sequence (P2), GGGACA/19 bp/TTTAAT (positions 657–687), just 9 bp upstream of the start site (S2) of the major gltA transcript (Fig. 4). This analysis also confirmed that promoter (P1), TTACAA/18 bp/TATAAT (positions 551–580), which is the same as promoter B suggested by Wood et al. (1984), is the most likely promoter for the minor transcript (Fig. 4). Its distance from the proposed start point (S1) is rather large (13 bp), and unless an atypical promoter is involved, it is possible that the true start is closer to the promoter, because the intervening region is very AT-rich and may have allowed terminal melting of a somewhat longer hybrid duplex.

**Location of the gltA transcription terminus.** Probes C and D, used to map the termination site (Fig. 1), generated S1-resistant hybrids of ~158 and ~980 bp, respectively (Fig. 2). These locate the transcriptional terminus very close to the end of the gltA gene at positions ~2224 (C) and
Transcripts of the gltA and sdhCDAB genes

Fig. 2. Transcript mapping of the gltA gene. Photographs of ethidium-bromide-stained non-denaturing polyacrylamide gels showing nuclease-S1-resistant RNA:DNA hybrids formed with probes A–D and A’–D’ (see Fig. 1). RNA was prepared from ED8641(pGS91) grown in L broth (−) or L broth plus glucose (+), and either 200 μg (a) or 400 μg (b) was added per track. The calibration lanes (S) contain the 123 bp ladder.

The size of the smaller band was found to be 148 bp when labelled samples were run alongside a sequence ladder on a denaturing gel (Fig. 3a). This accurate sizing locates the transcriptional terminus at position 2214 (Fig. 4). The terminus region contains a GC-rich sequence, capable of forming a stable stem-loop structure in the transcript, followed by a run of T(U) residues, which is characteristic of rho-independent terminators (Rosenberg & Court, 1979). The results confirm the earlier prediction of Ner et al. (1983), that this is the effective gltA terminator. No readthrough transcription into a distal gene(s) was detected; nor was there any evidence for any independent transcription in the downstream region. This region contains an unidentified open reading frame starting at position 2568, and another in the complementary strand ending at position 2556 that is fused (in phase) to the tet gene in pGS91 (Fig. 1). Assuming that transcription is continuous through the gltA gene then the sizes of the two transcripts initiated by the two gltA promoters, P₁ and P₂, are 1622 and 1519 bp, respectively.

Transcription analysis of the sdh genes

The sdh transcripts were analysed with RNA derived from ED8641(pGS65) as a potentially enriched source. This plasmid contains the 4.5 kb BamHI fragment (B₁–B₂) encoding the four sdh genes (sdhCDAB) inserted into the tet gene of the vector (Hull et al., 1983; Wood et al., 1984). It is presumed to contain the sdh promoter because the isolated BamHI fragment directs the synthesis of polypeptides corresponding to the four gene products (14 kDa, sdhC; 13 kDa, sdhD; 64 kDa, sdhA; 27 kDa, sdhB) in an in vitro transcription-translation system. Analogous studies with a 2.4 kb HindIII–XhoI sub-fragment have likewise indicated the presence of a second sdh promoter immediately upstream of the sdhD gene (R. J. Wilde, unpublished). The relevant segment of DNA and the M13 probes used in the transcript analysis are illustrated in Fig. 5. Seven different M13 probes (B', E, F, G, H, I and J) were used to map the sdh transcripts, and five complementary probes (B’, F’, G’, I’ and J’) were used as controls.

Location of sdh transcription start points. The S1-resistant hybrids formed with probes B' and E to I are shown in Figs 2 and 6. Probes B', E and F generated hybrids of ~400, ~225 and
Fig. 3. Autoradiograms of denaturing polyacrylamide gels for sizing the \(^{32}\text{P}\)-labelled DNA strands from nuclease-S1-resistant hybrid molecules. The sequence ladders (G, A, T, C) from probe A provide the calibration: distances from the first base of the sequencing primer are indicated. The bands obtained with probes (a) A and C, and (b) G and E are shown and the coordinates of the corresponding end-points are given in parentheses (see Figs 1 and 4, and 5 and 7).
corresponds to the full length of the probe (223-mer) plus one flanking nucleotide in the vector upstream of the sdhD structural gene (Fig. 7). The larger but weaker probe G product (224-mer) obtained with probe G locates the other start point and sizing the ladder (Fig. 6). The start points were defined more accurately by starting with radioactive probes which increases the extent of hybridization with the sdh transcript, only structural gene (Fig. 7). Likewise, the prominent band (128-mer) of the strong triplet (127 to 129-mer) corresponded to the intergenic region (S1, S2). Some of the probes also produced bands that are identifiable as the products of full-length transcripts of the sdhC gene showing (a) the proposed promoter sequences (P1, and P2; -35 and -10) and the transcriptional start points (S1 and S2) in the upstream region, and (b) the terminator in the downstream region. Also shown are four potential CRP sites, a ribosome binding site (RBS), regions of hyphenated dyad symmetry (underlined by converging arrows) and restriction sites (abbreviated as in Fig. 1).

Fig. 4. Nucleotide sequences flanking the gltA gene showing (a) the proposed promoter sequences (P1, and P2; -35 and -10) and the transcriptional start points (S1 and S2) in the upstream region, and (b) the terminator in the downstream region. Also shown are four potential CRP sites, a ribosome binding site (RBS), regions of hyphenated dyad symmetry (underlined by converging arrows) and restriction sites (abbreviated as in Fig. 1).

~ 660 bp (respectively), which correspond to a transcript with a start point (S1) in the gltA-sdhC intergenic region ~ 219 bp upstream of the sdhC gene. Two other hybrids generated by probes F (~ 220 bp) and G (~ 128 bp) correspond to a second and potentially major transcript having a start point (S2) which is located within the sdhC gene ~ 160 bp upstream of the sdhD gene (Fig. 5). Some of the other hybrids generated by probes F, G, H, and I, only B generated a detectable hybrid (~ 150 bp) which is derived from the gltA transcript (Figs 2 and 5).

The start points were defined more accurately by starting with radioactive probes (E and G) and sizing the DNA strands of the S1-resistant hybrids in denaturing gels alongside a sequence ladder (Fig. 3b). The smaller but more intense band of the doublet (225 to 226-mer) obtained with probe E locates the start point (S1) at position 872, exactly 219 bp upstream of the sdhC structural gene (Fig. 7). Likewise, the prominent band (128-mer) of the strong triplet (127 to 129-mer) obtained with probe G locates the other start point (S2) at position 1318, some 156 bp upstream of the sdhD structural gene (Fig. 7). The larger but weaker probe G product (224-mer) corresponds to the full length of the probe (223-mer) plus one flanking nucleotide in the vector which increases the extent of hybridization with the sdh transcript.
Fig. 5. Diagram showing the M13 probes used to map the *sdh* transcripts and segments of chromosomal and plasmid (pGS65 and pGS91) DNA presented as in Fig. 1; additional restriction sites are X (XhoI), and Sacl and NruI, which refer to *SalI* and *NruI* sites in the vector (pBR322). Left to right corresponds to clockwise in the *E. coli* linkage map. Coordinates are numbered according to Wood et al. (1984); position 1 is the first base of the *NruI* site and corresponds to coordinate 1264 in Fig. 1. The coordinates for specific probes are: 536-1274 (B, B'); 787-1096 (E); 726-1532 (F, F'); 1222-1444 (G, G'); 1455-1805 (H); 1692-2325 (I, I'); 3899-4615 (J, J'); and 3609-4992 (K, K').

Fig. 6. Transcript mapping of the *sdh* genes. Photographs of ethidium-bromide-stained non-denaturing polyacrylamide gels showing S1-resistant hybrids formed with probes E to I (see Fig. 5). RNA was prepared from ED8641(pGS65) and 200 μg was added per track. The calibration lanes (S) contain the 123 bp ladder.

A search for potential promoters revealed two high-scoring sequences (P1 and P2) associated with the *sdh* transcription start points (Fig. 7). These had been detected previously and designated D and E by Wood et al. (1984). The first, TTGTAA/17 bp/TATACT (P1: positions 835–863), is 9 bp upstream of *S*1 and presumably governs the transcription of all four *sdh* genes. The second, TGGGCA/16 bp/TATCAT (P2: positions 1284–1311), is 7 bp upstream of *S*2 and presumably directs the synthesis of an *sdhDAB* transcript.

**Location of the *sdh* transcriptional terminus.** The location of the *sdh* termination site was tentatively assigned to the 4364–4374 region during an earlier transcription analysis of the adjacent *suc* operon using the probes designated J, J', K and K' in Fig. 5 (Spencer & Guest,
Transcripts of the gltA and sdhCDAB genes

Fig. 7. Nucleotide sequences in the sdhCDAB region showing (a) the proposed sdh promoter sequences (P₁ and P₂; -35 and -10) and the corresponding start points (S₁ and S₂) in the gltA-sdhD segment, and (b) the termination site and REP sequences in the sdhB-sucA intergenic region. Other features are described in the legend of Fig. 4.

1985). The labelled DNA from the Sl-resistant hybrid produced with probe J has since been sized by the high resolution procedure using an extended sequence reaction of probe C' as the calibration ladder. It contained a cluster of four oligonucleotides (465 to 468-mers) the strongest being the 466-mer which corresponds to a major termination site at position 4364 (data not shown). There is no typical rho-independent terminator, but the site immediately follows the first of three highly conserved REP sequences in the sdhB-sucA intergenic region where there is a sequence, CAACCAG (positions 4358-4364), which resembles one that is found at the 3'-terminus of some rho-dependent terminators (CAATCAA; Lau et al., 1984). Most REP
sequences are transcribed (Higgins & Smith, 1986) but no evidence was found for readthrough transcription across the other REP sequences in the sdhB–sucA intergenic region, though independent transcription from two suc promoters was detected (Fig. 5).

The probes include all but the distal segment of the sdhA gene and assuming that transcription is continuous, the sdh promoters generate two polycistronic mRNAs, a long 3493 base sdhCDAB transcript (P1) and a shorter 3047 base sdhDAB transcript (P2).

**Effect of glucose on gltA and sdhCDAB transcription**

The existence of two promoters for the gltA and sdh genes suggested that they may provide a means for differential expression of citrate synthase and succinate dehydrogenase. It is known that the synthesis of these enzymes is subject to repression by glucose during expression both from the chromosome and from multicopy plasmids (Bloxham et al., 1983; Robinson et al., 1983; R. J. Wilde & J. R. Guest, unpublished). This repression is presumably mediated by 3'–5' cyclic AMP (cAMP) and the cAMP receptor protein (CRP), which form a cAMP–CRP complex that activates transcription by interacting with the CRP-binding sites in the promoter regions of the corresponding genes. In the presence of a preferred substrate such as glucose, the cAMP concentration falls so that the transcription of catabolite-sensitive genes and operons can no longer be activated.

Studies with several samples of RNA from cultures of ED8641(pGS91) grown in L broth with added glucose (1%, w/v) and probe A produced no detectable RNA:DNA hybrids (Fig. 2). It would thus appear that transcription from both the major (P2) and minor (P1) gltA promoters is subject to catabolite repression. Further experiments using radioactive probes confirmed that the synthesis of both transcripts is repressed, although the detection of small amounts of the longer transcript was impaired by the presence of some non-specific labelled material in the corresponding region of the gels. In comparable studies with samples of RNA from ED8641(pGS65) and probe F (32P-labelled and unlabelled) no S1-resistant hybrids corresponding to either sdh transcript could be detected following growth in the presence of glucose (data not shown). It is concluded that transcription from both sdh promoters is subject to catabolite repression.

Potential CRP-binding sites were sought in both strands of the relevant gltA–sdhD region using the consensus sequence of Chapon & Kolb (1983) (see Methods). Two sequences that are identical at 12 out of 17 positions are shown as CRP1 and CRP2 in Figs 4 and 7. Two further sequences (CRP3 and CRP4) that are as homologous in the two internal 5 bp segments, but have less overall homology (9/17), are also shown in Fig. 4. The sdhC gene contains just one comparable sequence, which is designated CRP5 in Fig. 7. There are clearly potential CRP sites overlapping or upstream of the gltA promoters, P1 (CRP4) and P2 (CRP1 and CRP3), and the sdh promoters, P1 (CRP2) and P2 (CRP5). However, it is not known whether these are functional sites, nor whether the CRP2 site could perform a dual role in activating the P1 promoters of both the sdh and gltA genes.

**DISCUSSION**

The transcript analysis has identified two diverging pairs of transcripts in the gltA–sdhCDAB region. The major gltA transcript (1519 bases; promoter P2) and the minor transcript (1622 bases; promoter P1) start 196 and 299 bp upstream of the gltA structured gene and terminate 39 bp downstream of the translational stop codon. The gltA transcripts appeared to be monocistronic, but the transcription terminus is unusual in being within the loop of the potential stem-loop structure of a typical rho-independent terminator, rather than beyond the stem. This is a reproducible feature and it is not known whether it is due to post-transcriptional processing or degradation, interference from converging (but undetected) transcription, or an artefact of the method. The major sdhDAB transcript (3047 bases; promoter P2) and minor sdhCDAB transcript (3493 bases; promoter P1) start 156 and 219 bp upstream of the sdhD and sdhC structural genes, respectively, and terminate 45 bp downstream of the translational stop codon. No readthrough transcription into the adjacent sucABCD operon was detected, but nor could the
effects of post-transcriptional processing or the influence of the REP sequences on transcript stability be assessed. Nevertheless, the results offered no support for the attractive possibility that the sdh and suc genes, encoding the three citric acid cycle enzymes that are most severely repressed under anaerobic conditions (2-oxoglutarate dehydrogenase complex, succinyl-CoA synthetase and succinate dehydrogenase), are co-transcribed and co-regulated at a common promoter-operator site. It is apparent in both cases that the shorter transcripts are more abundant than their longer counterparts, but the reason for this is not clear.

The promoters were chosen primarily for their location and homology with the E. coli consensus. However, because the sequences of positively controlled promoters (Raibaud & Schwartz, 1984) and those involving minor sigma factors, such as the htpR and ntrA gene products (Grossman et al., 1984; Hirschman et al., 1985), deviate from the consensus, the operation of alternative promoters cannot be ruled out. Interestingly, when compared with the 19 base consensus, the promoters for the major (shorter) transcripts (P2) are less homologous than those for the minor transcripts (P1). This could mean that transcription from the P2 promoters is controlled by a gene-activator protein(s). Furthermore, the presence of two functional promoters could provide mechanisms for the differential regulation of enzyme synthesis. In analogous cases, e.g. lac, gal and deo operons, in vitro studies have shown that one promoter requires the cAMP–CRP complex for activation (CRP-dependent), whereas the other does not (Spasskey et al., 1984; Valentin-Hansen et al., 1985). In each case the CRP-independent promoter generates the longer transcript but there is no consistency about the positions of the CRP-binding sites, which are overlapping and/or upstream of the corresponding −35 regions. All of the gltA and sdh promoters are associated with putative CRP-binding sites, though none possess the perfect 5 bp motif, and all have −35 sequences that fall within the generally poor range of homology observed for CRP-dependent promoters (Busby, 1986; Raibaud & Schwartz, 1984). None of the gltA or sdh transcripts could be detected in glucose-grown organisms. For the gltA gene, this suggests that the attractive possibility that P2 is CRP-dependent whereas P1 is CRP-independent and maintains a low level of citrate synthase during glucose repression is untenable. If stable, the amount of gltA transcript produced during glucose repression is presumably too small to be detected by the methods applied here. In the case of succinate dehydrogenase, no sdh transcription would be expected during glucose repression, rather some derepression of fumarate reductase occurs. However, the fact that the major transcript lacks message for the cytochrome subunit (sdhC gene product) poses other problems concerning the subunit stoichiometry of the assembled complex. Unless the cytochrome is overproduced from the minor transcript, the complex must possess a disproportionately low content of this subunit. That this may be the case is apparent from the relative amounts of the sdhC gene product detected during in vitro transcription–translation studies (Wood et al., 1984; R. J. Wilde, unpublished). Disproportionate subunit ratios such as those in the ATPase complex (Walker et al., 1984) may be found for the hydrophobic subunits and this could presumably be accommodated by differential expression at both the transcriptional and translational levels. Despite their proximity, the gltA and sdh promoters would appear capable of functioning independently, though some degree of interaction would seem likely. Clearly more work, including in vitro transcription and footprinting studies, will be needed to identify the CRP-binding sites and to define the mechanisms whereby diverse environmental factors interact in controlling the expression of these genes. It will also be interesting to identify the transcriptional features that are shared with other citric acid cycle genes that respond to the same regulatory stimuli.

We thank Dr Margaret E. Spencer for valuable discussions. This work was supported by the Science and Engineering Research Council.

REFERENCES


Transcripts of the gltA and sdhCDAB genes


WEITZMAN, P. D. J. (1981). Unity and diversity in some bacterial citric acid-cycle enzymes. Advances in Microbial Physiology 22, 185–244.