Biochem. J. (1989) 260, 737–747 (Printed in Great Britain)

Overexpression and site-directed mutagenesis of the succinyl-CoA synthetase of *Escherichia coli* and nucleotide sequence of a gene (*g30*) that is adjacent to the *suc* operon

David BUCK* and John R. GUEST†
Department of Microbiology, University of Sheffield, Western Bank, Sheffield, S10 2TN, U.K.

The succinyl-CoA synthetase of *Escherichia coli* is encoded by two genes, *sucC* (β subunit) and *sucD* (α subunit), which are distal genes in the *sucABCD* operon. They are expressed from the *suc* promoter, which also expresses the dehydrogenase and dihydrolipoyl succinyl-transferase subunits of the 2-oxoglutarate dehydrogenase complex. Strategies have now been devised for the site-directed mutagenesis and independent expression of the succinyl-CoA synthetase ($\alpha_2\beta_2$ tetramer) and the individual subunits. These involve (1) subcloning a promoterless *sucCD* fragment downstream of the *lac* promoter in M13mp10, and (2) precise splicing of the *suc* coding regions with the efficient *atpE* ribosome-binding site and expression from the thermoinducible λ promoters in the pJLA503 vector. Succinyl-CoA synthetase specific activities were amplified 40–60-fold within 5 h of thermoinduction of the λ promoters, and the α and β subunits accounted for almost 30% of the protein in supernatant fractions of the cell-free extracts. Site-directed mutagenesis of potential CoA binding-site residues indicated that Trp-43β and His-50β are essential residues in the β-subunit, whereas Cys-47β could be replaced by serine without inactivating the enzyme. No activity was detected after the histidine residue at the phosphorylation site of the α-subunit was replaced by aspartate (His-246α → Asp), but this alteration seemed to have a deleterious effect on the accumulation of the enzyme in cell-free supernatant extracts. The nucleotide sequence of an unidentified gene (*g30*) that is adjacent to the *sucABCD* operon was defined by extending the sequence of the citric acid cycle gene cluster by 818 bp to 13379 bp: *gltA-sdhCDAB-sucABCD-g30*. This gene converges on the *suc* operon and encodes a product (P30) that contains 230 amino acids (M_r 27251). Highly significant similarities were detected between the *N*-terminal region of P30 and those of GENA [the product of another unidentified gene (*gene*A) located upstream of the *aceEF-lpd* operon], and GNTR (a putative transcriptional repressor of the gluconate operon of *Bacillus subtilis*). Possible roles for GENA and P30 as transcriptional regulators of the adjacent operons encoding the pyruvate and 2-oxoglutarate dehydrogenase complexes are discussed.

INTRODUCTION

The succinyl-CoA synthetase [SCS; succinate:CoA ligase (ADP-forming); EC 6.2.1.5] of *Escherichia coli* catalyses the reversible thioesterification of succinate in three steps involving the phosphoryl enzyme and enzyme-bound succinyl phosphate intermediates:

$$\text{Succinate} + \text{CoA} + \text{ATP} \overset{Mg^{2+}}{\rightleftharpoons} \text{Succinyl-CoA} + \text{ADP} + P_i.$$

It functions in the citric acid cycle, providing an important site of substrate level phosphorylation, and a source of succinyl-CoA for anabolic purposes when the route from 2-oxoglutarate is repressed, e.g. during anaerobiosis. The enzyme is assembled as a non-dissociating tetramer ($\alpha_2\beta_2$) from two types of subunit: the α-subunit (M_r 29600), which contains the histidine residue that is phosphorylated during catalysis and the ATP binding site; and the β-subunit (M_r 41400), which contains the binding sites for succinate and CoA (Bridger, 1974; Nishimura, 1986). The corresponding enzymes from eukaryotes and Gram-positive bacteria differ in being non-associating dimers ($\alpha\beta$) with different nucleotide specificities (Weitzman, 1981; Wolodko *et al.*, 1986). Chemical modification studies (with the *E. coli* enzyme) have indicated important functional roles for two cysteine residues, a tryptophan residue and one histidine residue in addition to the phosphorylation site (Nishimura, 1986), and the enzyme has been crystallized for structure determination (Wolodko *et al.*, 1984).

The two subunits of the SCS of *E. coli* are encoded by adjacent genes, *sucC*(β) and *sucD*(α), at the distal end of the *sucABCD* operon, which also encodes the specific dehydrogenase (E1o, *sucA*) and succinyltransferase (E2o, *sucB*) components of the 2-oxoglutarate dehydrogenase complex (see Fig. 1; Darlison *et al.*, 1984; Spencer *et al.*, 1984; Buck *et al.*, 1985, 1986). The *suc* operon forms part of a major cluster of nine citric acid cycle genes specifying four enzymes, including citrate synthase (*gltA*) and the succinate dehydrogenase complex (*sdh*): *gltA-sdhCDAB-sucABCD* (Miles & Guest, 1987). This gene cluster is located at 16.8 min in the *E. coli* linkage map and it is included in a fully-sequenced

Abbreviations used: SCS, succinyl-CoA synthetase; PDH, pyruvate dehydrogenase; IPTG isopropyl-β-D-thio-galactoside; PAGE, polyacrylamide-gel electrophoresis.

* Present address: Schering Agrochemicals Ltd., Chesterford Park Research Station, Saffron Walden, Essex CB10 1XL, U.K.
† To whom correspondence should be addressed.

13061 bp segment of DNA (Buck *et al.*, 1985). A transcript analysis has indicated that SCS is expressed from a readthrough transcript (*sucABCD*), which is the longer of the two transcripts (*sucAB* and *sucABCD*) that encode the E1o and E2o components of the 2-oxoglutarate dehydrogenase complex (Fig. 1; Spencer & Guest, 1985). This is consistent with the absence of a recognizable promoter upstream of the *sucC* gene and evidence for co-ordinate regulation of the synthesis of SCS and the 2-oxoglutarate dehydrogenase complex (Buck *et al.*, 1985, 1986). The nucleotide sequence has likewise shown that the stop codon of *sucC* overlaps the initiation codon of *sucD* by a single base pair (Buck *et al.*, 1985), and this could provide a mechanism for translational coupling of the β and α subunits, which are known to be maintained in strict equimolar proportions (Wolodko *et al.*, 1980). The sequence of the α-subunit of rat liver mitochondrial SCS has recently been deduced from a cDNA sequence and it exhibits a remarkable 70% sequence identity with the corresponding *E. coli* subunit, despite the differences in structural organization of the two enzymes (Henning *et al.*, 1988).

In order to facilitate further studies on the structural and mechanistic aspects of SCS, the *sucCD* genes were subcloned into appropriate vectors to provide a useful strategy for over-expression and site-directed mutagenesis of the enzyme. Several mutants with alterations at potential active-site residues were constructed, and the nucleotide sequence of an unidentified gene (*g30*) situated adjacent to the *suc* operon and encoding a 30kDa product (P30) was also defined.

MATERIALS AND METHODS

E. coli strains, plasmids and bacteriophages

The following *E. coli* strains were used as hosts for M13 and plasmid derivatives: JM101 (Δ*lac-proAB supE* thi/F′*traD36 proA*$^+$*B*$^+$ *lacI*^{q}*ZΔM15*) for the routine preparation of DNA templates for sequence analysis (Messing, 1979); RZ1032 (HfrKL16, *dut-1 ung-1 thi-1 relA1 spoT1 zbd-279*::Tn*10 supE44*) for producing uracil-containing DNA templates for mutagenesis (Kunkel, 1985); BMH71-18mutL (Δ*lac-proAB supE thi mutL*::Tn*10*/F-′*traD36 proA*$^+$*B*$^+$ *lacI*^{q}*ZΔM15*) used for increasing the mutation efficiency (Kramer *et al.*, 1984); and HB101 (F$^-$, *recA13 proA2 recA13* thi-*1 ara-14 lacY1 galK2 xyl-5 mtl-1 rpsL20 supE44 hsdS20*) as a transformation recipient (Boyer & Roulland-Dussoix, 1969). Plasmids pGS130, pGS131 and pGS132 (Buck *et al.*, 1986) were used as sources of the *sucCD* and *g30* genes; they were prepared in a Dam$^-$ strain so as to be susceptible to *BclI* digestion. The $\lambda P_R \lambda P_L$-expression vector pJLA503 (Schauder *et al.*, 1987) was kindly provided by Dr. J. E. G. McCarthy, GBF-Gesellschaft für Biotechnologische forschung, Braunschweig, Germany. The M13mp8-19 series of phages (Messing & Vieira, 1982; Norrander *et al.*, 1983; Yanisch-Perron *et al.*, 1985) were used in nucleotide sequence analysis and mutagenesis.

Molecular cloning

Standard methods were used for plasmid preparation, restriction endonuclease digestion, isolation of DNA fragments, ligation, transformation and transfection (Maniatis *et al.*, 1982). The *sucC* and/or *sucD* genes were transferred from M13 to the expression vector (pJLA503) by primer-extension cloning from single-stranded M13 DNA (Miles & Guest, 1985). The double-stranded product was restricted with *Nde*I and *Bam*HI or *Sal*I and the corresponding fragment(s) was cloned into appropriately cut and phosphatased pJLA503.

Cloning in M13 and nucleotide sequence analysis

The nucleotide sequence of the unidentified gene (*g30*) was determined by extending the sequence of the previously analysed *sucCD* region (Buck *et al.*, 1985). A combination of 'shotgun' cloning and 'directed' cloning of specific fragments derived from the 2.5 kb *Nru*I (N_5–N_6) and 1.3 kb *Sal*I–*Bcl*I (Sa_2–Bc_3) fragments of pGS131 were used (see Fig. 5 legend for definition of abbreviations). Some additional sequence data were obtained from the 1.2 kb *PvuI*–*Eco*RV$_{pBR322}$ fragment of pGS130 (see Fig. 5). Specific fragments (particularly C_1–C_2, Sa_2–Bc_3, C_3–Bc_3) and randomly-generated subfragments (*Sau*3A and *Msp*I) were cloned into appropriate sites of different M13 vectors, and single-stranded DNA templates were prepared and sequenced by the dideoxy chain-termination method using 'universal' primer, [α-^{35}S]thio-dATP, and buffer-gradient gels (Sanger *et al.*, 1980; Biggin *et al.*, 1983). Nucleotide sequences were compiled and analysed with the aid of the Staden computer programs (Staden, 1979, 1980; Staden & McLachlan, 1982). Sequence comparisons with DNA and protein databanks were done with the EMBLSCAN and PEPSCAN programs (Bishop & Thompson, 1984; Bishop, 1984), or by Dr. J. F. Collins of the Biocomputing Research Unit, Department of Molecular Biology, Edinburgh University, Edinburgh, Scotland, using the Dap Protein Database Search (Coulson *et al.*, 1987; Collins *et al.*, 1988).

Oligonucleotide-directed mutagenesis

The 'two-primer' method was performed according to Zoller & Smith (1984) except that uracil-containing M13 templates were used (Kunkel, 1985). Mutant strands were selected by transfection of strain BMH71-18mutL (*ung*$^+$) and M13 plaque formation in a lawn of *E. coli* JM101 to prevent unnecessary exposure of the phages to the mutator strain. The dot-hybridization procedure of Zoller & Smith (1983) involving ^{32}P-labelled mutagenic primer as the probe was used to identify mutant phages, but only where stated. Synthetic 'universal' primer (17-mer) was used in conjunction with several additional synthetic oligonucleotides for directing the mutations and for confirming the sequences of the products. These primers, including their binding co-ordinates in the nucleotide sequence of Buck *et al.* (1985) and their functions, were as follows: S27, ACCCGCGT^{C}GCCCATACG (6543–6526; His-246α → Asp); S47, AAGTTCAT^{A}TGTTCTGT (4638–4622; *Nde*I site in *sucC*); S48, AAATGGACAT^{ATG}TTCCCCTCCACT (5806–5782, *Nde*I site in *sucD*); S49, TTTCACTACC^{AC}-CGGACCGGCAC (4768–4746; Trp-43β → Val); S50, GTGAACCTGA^{G}ATTTCACTAC (4780–4760; Cys-47β → Ser and sequencing); S51, GGCCACCAGC^{TGC}-AACCTGACAT 4790–4768; His-50β → Ala); S52, TTTAGCGATATCGGTCGC (4948–4931; specific sequencing primer); S53, GCCTGCGTCGATGGC (6063–6049, specific sequencing primer). They were designed to give strong and unique hybridization at the desired sites in the template DNAs.

Overexpression of SCS

The SCS activities were assayed in 50 ml cultures of JM101 infected with M13 derivatives after growth at 37 °C for 15 h in 2TY medium (Miller, 1972) ± isopropyl-β-D-thio-galactoside (IPTG) (200 μg/ml) following inoculation (1:100) from an overnight bacterial culture and 10^{11} plaque-forming units (p.f.u.) of the phage. The cells were harvested, washed in 40 mM-phosphate buffer (pH 7.8), resuspended in one-tenth of their original volume and disrupted by ultrasonic treatment. The cell-free supernatants obtained by centrifuging (see below) were assayed for total protein (Lowry *et al.*, 1951) and SCS activity (Bridger *et al.*, 1969), and they were also analysed by SDS/polyacrylamide-gel electrophoresis (SDS/PAGE).

The synthesis of SCS in HB101 transformants containing the expression plasmids pJLA503 and pGS202 was assayed using 70 ml shaken cultures grown at 28 °C in L broth (Lennox, 1955) plus ampicillin (100 μg/ml) starting with an inoculum (1:50) from an overnight culture. The cultures were shifted to 42 °C at A_{650} = 0.4–0.5, and samples (10 ml) were harvested at 0, 1, 2, 3 and 5 h. The cells were washed and disrupted as above, and sedimented at 16000 rev/min (30000 *g*) for 20 min. Samples of the supernatant and pellet fractions, as well as the unfractionated extract, were assayed for total protein and SCS activity, and were also analysed by SDS/PAGE.

SDS/PAGE

SDS/PAGE (0.1% SDS) was performed by the method of Laemmli (1970) using a 10% running gel with a 5% stacking gel. Each track contained 15 μg of protein and some gels were subjected to a quantitative densitometric analysis following staining with Coomassie Brilliant Blue.

Materials

The [α-^{35}S]thio-dATP (410 Ci/mmol) and [γ-^{32}P]ATP (3000 Ci/mmol) were supplied by New England Nuclear. Restriction endonucleases, DNA polymerase (Klenow fragment), T_4 DNA ligase and T_4 polynucleotide kinase were from Bethesda Research Laboratories or the Boehringer Corporation. The M13mp8-19 replicative-form DNAs were from Pharmacia-PL Biochemicals. The 'universal' sequencing primer was from Celltech and other oligonucleotides were supplied by courtesy of G. D. Searle Ltd.

RESULTS AND DISCUSSION

Strategy for mutagenesis and overexpression

An important prerequisite for a site-directed mutagenesis project is to be able to express the mutated genes at a high level either directly from the mutagenized DNA (phage or plasmid) or after facile transfer of a mutagenized cassette to a suitable expression vector. Both approaches have been investigated with the SCS genes. In one, M13mp10 derivatives were used for both mutagenesis and expression, and in the other, mutagenesis in M13mp10 was combined with expression from the thermoinducible $\lambda P_R P_L$ promoters of pJLA503 (Schauder *et al.*, 1987).

Overexpression from the *lac* promoter of M13mp10. Inspection of the nucleotide sequence of the *suc* region showed that the *sucC* and *D* structural genes are conveniently located within a 2.5 kb *Nru*I fragment within a minimum amount of flanking DNA (N_5–N_6 in Fig. 1). This fragment was isolated from *Nru*I-digested pGS132 and cloned into *Sma*I-cut and phosphatased M13mp10 RF. Two types of product differing in the orientation of the cloned fragment relative to *lac* promoter in the vector, were recovered and distinguished by sequence analysis. Representatives of each class (Φ1 and Φ2; Fig. 1) were tested for their ability to express SCS in the presence or absence of IPTG, and the results confirmed that the *suc* genes are expressed from the *lac* promoter in Φ1 (Table 1). There was also a very significant amplification of two polypeptides corresponding to the α and β subunits of SCS in the extracts of cells analysed by SDS/PAGE (results not shown). The highest specific activities obtained in the presence of the inducer (106 μmol/h per mg of protein) were greater than those obtained with pGS131 (14–25 μmol/h per mg of protein), where the cloned *sucC* and *D* genes are thought to be expressed from the *tet* promoter of the vector, pBR322 (Buck *et al.*, 1986). Clearly Φ1 is a very versatile derivative of M13mp10 that can be used for expressing SCS from the *lac* promoter, for mutagenesis *in vitro*, and for subcloning purposes because it retains most of the restriction sites in the vector polylinker.

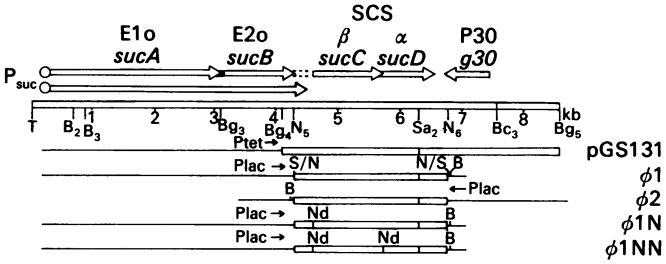

Fig. 1. Map of the *sucABCD* operon of *E. coli* and subcloning of the SCS (*sucC* and *D*)

The positions of the *suc* genes encoding the specific E1o and E2o components of the 2-oxoglutarate dehydrogenase and the α and β subunits of SCS are aligned with the corresponding restriction map. The position of an unidentified gene (*g30*) is also indicated. The restriction sites for *Bam*HI(B), *Bcl*I(Bc), *Bgl*II(Bg), *Nru*I(N), *Sal*I(Sa) and *Sma*I(S), are numbered according to Buck *et al.* (1985) and the nucleotide positions are numbered in kb starting at an arbitrary *Taq*I site (T). The mRNA transcripts are indicated as are relevant promoters in the *E. coli* and vector DNAs. Plasmid pGS131 contains the 4.5 kb *Bgl*II fragment cloned into the *Bam*HI site of pBR322 (pGS132 contains the same fragment in the opposite orientation). The M13mp10 derivatives (Φ1–4) contain the 2.5 kb *Nru*I fragment cloned into the vector *Sma*I site and the positions of the newly-created *Nde*I sites (Nd) are indicated.

Overexpression from thermoinducible λ promoters. In order to achieve even higher specific activities, the *sucC* and *D* genes were cloned into pJLA503 (Fig. 2). This vector contains an *Nde*I site (CAT*ATG*) that facilitates precise cloning of coding regions at an optimal distance from the highly-efficient *atpE* translational initiator, and once cloned their expression is controlled by tandem λP_R

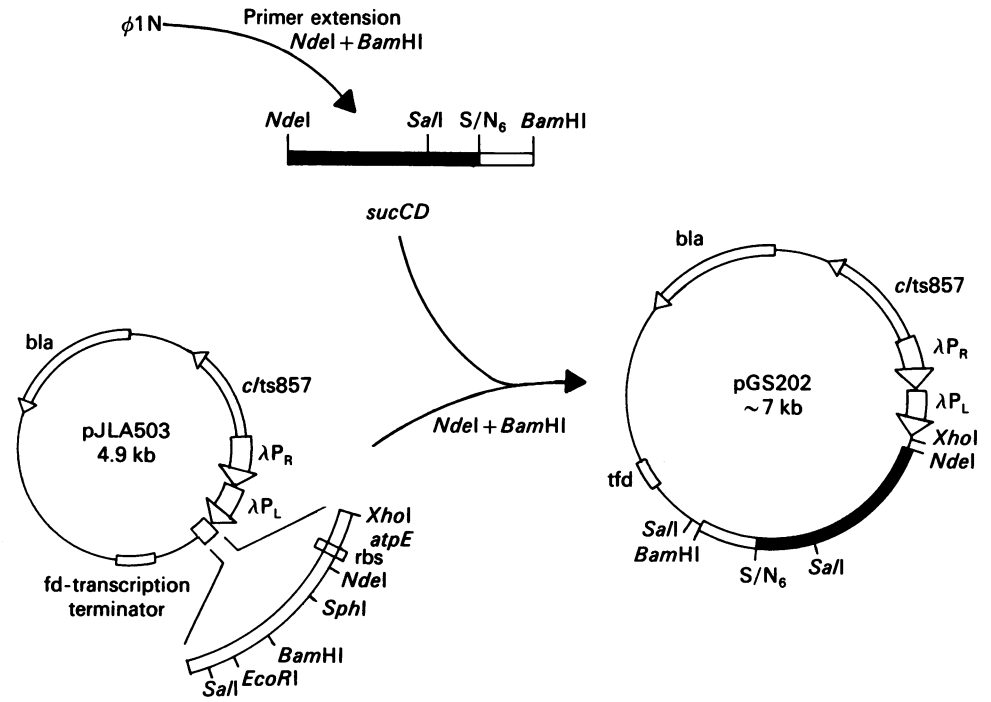

Fig. 2. Strategy for cloning the *sucC* and *D* genes into the expression vector pJLA503 to generate pGS202

The expression vector pJLA503 contains tandem λ promoters controlled by the thermosensitive λ*cI*857 repressor and located upstream of the *atpE* translation initiator region and cloning site (99 bp expanded region). The 2.2 kb *Nde*I–*Bam*HI fragment (*sucC* and *D*) was obtained from primer-extended Φ1N (Fig. 1) and cloned into the corresponding sites of pJLA503 to generate the SCS expression plasmid.

and P_L promoters, which are repressed at 28 °C but induced at 42 °C, due to inactivation of the λ*cI*857 thermosensitive repressor (Schauder *et al.*, 1987). In order to use pJLA503 for expressing SCS, an *Nde*I site had to be created at the initiation codon of the *sucC* gene. The chosen vector has the advantage over related vectors containing *Nco*I (CC*ATG*G) or *Sph*I (GC*ATG*C) sites because *Nde*I sites can be introduced without affecting the codon adjacent to the formylmethionine codon, even though *Nde*I sites proved to be relatively resistant to cleavage by the corresponding enzyme. Uracil-containing single-stranded DNA of Φ1 was used as the template for generating the *Nde*I site by the 'two-primer' method with the 'universal' sequencing primer and a mutagenic primer, S47 (see the Materials and methods section). The mutagenic primer hybridizes to positions 4638–4622 in *sucC* (Buck *et al.*, 1985) except for a single mismatch to generate the C → T mutation at position 4630 (Fig. 3). Of 24 potential mutant phages tested by dot-blot hybridization with ^{32}P-labelled S47 as the probe, 16 gave positive signals both at room temperature and at 44 °C. Five of these were purified and sequence analyses with S50 as primer (see the Materials and methods section) confirmed the presence of the desired *Nde*I site in four of the isolates. A representative phage was designated Φ1N. The fifth isolate contained an insertion of 10 bp (A^T*ATG*AACTT) which generated a tandem duplication that included the C → T mismatches and the translational initiation codon. This could have arisen because the mutagenic primer and the 5′ end of the *sucC* coding region both contain direct repeats and some internal complementarity, which could stabilize the binding and mispairing of two primer molecules in the critical region. The corresponding mutant was not investigated further.

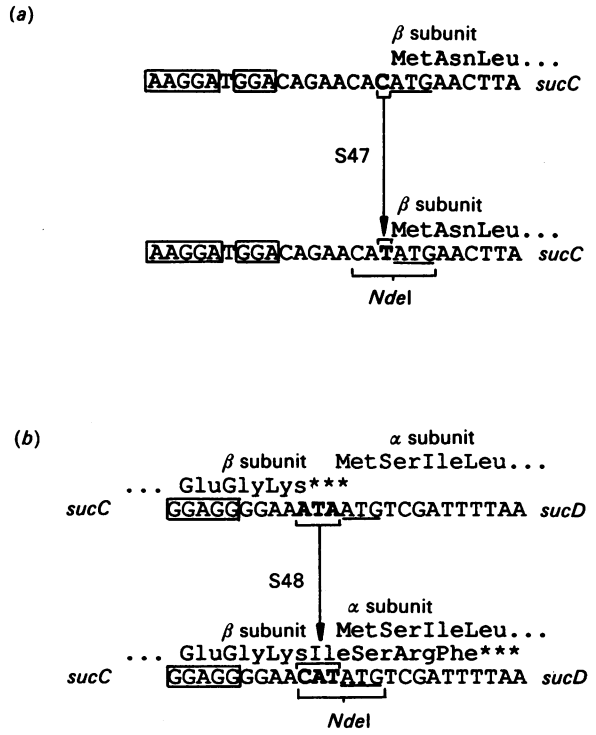

Fig. 3. Introduction of *Nde*I sites at the 5′ ends of the *sucC* and *sucD* coding regions

The sequence changes associated with introducing *Nde*I sites at the translational initiation codons of (*a*) the *sucC* gene and (*b*) the *sucD* gene. The ribosome-binding sites are boxed, start codons are underlined, and stop codons are denoted by three asterisks.

Table 1. Overexpression of SCS from *lac* and λ promoters with phage (M13) and plasmid (pJLA503) vectors

Cultures of JM101 were infected with M13mp10 derivatives containing the *sucCD* region in two orientations, and grown for 15 h with or without IPTG (200 μg/ml); see the Materials and methods section. Transformants of HB101 containing the expression vector pJLA503 or its derivative pGS202 were grown for 5 h at 28 °C and then for a further 5 h at 42 °C. SCS activities were assayed in the supernatant fractions of ultrasonic extracts, and are expressed as μmol of substrate transformed/h per mg of protein.

Source of *sucC* and *D* genes	Inducer	Specificity activity of SDS (μmol/h per mg of protein)	
		Uninduced	Induced
M13mp10(vector)	IPTG	11	10
Φ1	IPTG	15	106
Φ2	IPTG	7	10
Φ1N	IPTG	13	112
pJLA503(vector)	42 °C	11	15
pGS202	42 °C	14	617

The presence of the *Nde*I site in Φ1N had no effect on the IPTG-inducible expression of SCS (Table 1) so the *sucCD* region was subcloned into pJLA503 as a 2.2 kb *Nde*I–*Bam*HI fragment. This was derived from Φ1N by primer extension and endonuclease digestion and the structure of the resulting plasmid, pGS202 (Fig. 2), was confirmed by restriction analysis. As anticipated, the SCS specific activities of thermally-induced pGS202 transformants were amplified some 40–60-fold, which is considerably higher than obtained previously with pBR322 and M13 derivatives (Tables 1 and 2). Inclusion bodies appeared during overexpression of the *sucCD* region with pGS202 so it was of interest to investigate the progress of enzyme synthesis and its distribution in the supernatant and pellet fractions of crude ultrasonic extracts. The results of a typical experiment are shown in Table 2. The synthesis of SCS increased rapidly during the first 2 h at 42 °C and there were parallel increases in specific activity. Most of the activity (> 90 %) was located in the supernatant fraction, which in turn contained 50–65 % of the total protein in the culture. The course of expression was analysed by SDS/PAGE and staining of the proteins in the supernatant and pellet fractions of the cell extracts (Fig. 4). The amounts of the α and β subunits relative to the total protein in each fraction were estimated by densitometric analysis of the stained gels (Table 2). This indicated that as much as 24–29 % and 17–20 % of the protein in the respective supernatant and pellet fractions consisted of subunits of SCS. Assuming that the stain reacts equally with both types of subunit, and that the subunits are present in equimolar amounts, a ratio of α-to-β protein of 0.71 would be expected. This is true for the supernatant fractions, where the ratios range from 0.69–0.73, but there is a very significant excess of β subunits (or a shortfall of α subunits) in the pellet fractions, where the ratios are 0.25–0.45. This disproportionate synthesis suggests that the tight translational coupling between *sucC* and *sucD* may fail at high rates of expression, possibly due to a decline in the efficiency of translational reinitiation at the *sucD* (α) ribosome binding site. Alternatively, a significant proportion of potential *sucCD* transcripts may lack the *sucD* (α) segment, due to premature termination of transcription, or to incomplete mRNA processing from the 3′ end by the mechanism that may normally be responsible for generating *sucAB* as well as *sucABCD* transcripts. It would appear that the supernatant fractions contain subunits assembled into native tetramers ($\alpha_2\beta_2$), whereas the overproduced β subunits are consigned to the pellet fraction. It is not known whether the extra β subunits form a discrete

Table 2. Thermoinduction of SCS and distribution of enzyme activity and protein subunits in supernatant and pellet fractions of ultrasonic extracts

Strain HB101(pGS202) was grown in L broth at 28 °C before shifting to 42 °C at time zero. Cells were harvested from 10 ml samples and ultrasonic extracts were prepared, fractionated and assayed for SCS activity (see the Materials and methods section). The specific activities are expressed as μmol/h per mg of protein. The total activities (μmol/h) and protein (mg) refer to the amounts present in each fraction per mol of original culture, and the distribution of the total activity between the supernatant and cell fractions is also indicated as a percentage (in parentheses). The specific activities of SCS in control cultures of HB101(pJLA503) increased from 12 to 20 under comparable conditions. The relative amounts of α and β subunits expressed as percentages of the protein in the individual supernatant or pellet fractions were estimated densitometrically with stained gels after SDS/PAGE analysis (see Fig. 4).

Time at 42 °C (h)	Culture density (A_{650})	Supernatant fraction					Pellet fraction				
		Total SCS activity (μmol/h)	Protein (mg)	Specific activity (μmol/h per mg of protein)	α(%)	β(%)	Total SCS activity (μmol/h)	Protein (mg)	Specific activity (μmol/h per mg of protein)	α(%)	β(%)
0	0.40	2 (100)	0.05	38	–	–	0	0.03	0	–	–
1	0.59	58 (91)	0.07	860	10	14	6 (9)	0.07	88	5	11
2	0.70	119 (92)	0.10	1280	11	16	11 (8)	0.07	148	5	12
3	0.72	133 (92)	0.09	1450	11	15	12 (8)	0.07	167	3	14
5	0.72	170 (94)	0.10	1670	12	17	11 (6)	0.06	184	4	16

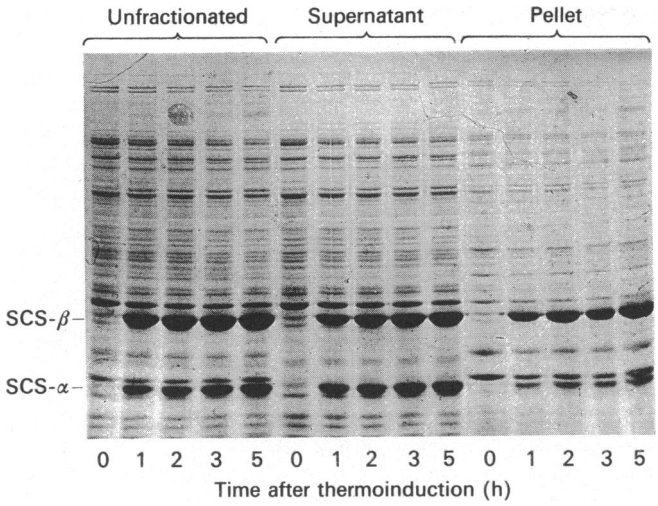

Fig. 4. Overexpression of SCS

The synthesis and distribution of the SCS α and β subunits were analysed in the unfractionated, supernatant and pellet fractions of ultrasonic extracts prepared at different times after thermoinduction of a pGS202 transformant of *E. coli* HB101. Equivalent samples of protein (15 μg) were separated by SDS/PAGE (10%) followed by staining with Coomassie Brilliant Blue. No polypeptides migrating to the same positions as the α and β subunits were amplified in analogous samples of a strain containing the vector, pJLA503.

sedimentable aggregate or whether they are incorporated into inclusions containing heterogeneous mixtures of both types of subunit. The disproportionate production of β subunits may indeed be responsible for inducing the formation of the inclusions. Relative to the supernatant fraction, the pellet contains considerably more α subunit than is reflected in its enzyme activity. This suggests that much of the α subunit is assembled into inactive aggregates, or into active but inaccessible complexes, in the inclusions. Based on its α content, the SCS activity of the pellet fraction should be about 3-fold greater than observed, i.e. about 20% of the enzyme in the induced cultures.

A strategy for overexpression of individual subunits. The strategy devised for the successful overexpression of both subunits of SCS offers the possibility of overexpressing each subunit independently. In the case of the β subunit an expression plasmid for the wild-type *sucD* gene could be constructed by a facile deletion of the 0.47 kb *Sal*I fragment of pGS202 *in vitro* (by *Sal*I digestion and religation, see Fig. 2). Alternatively, the 1.68 kb *Nde*I–*Sal*I fragment of Φ1N, and mutant derivatives of Φ1N (Fig. 1), could be subcloned between the corresponding sites of pJLA503, as was done for the *Nde*I–*Bam*I fragment containing the intact *sucC* and *sucD* genes (Fig. 2).

In the case of the α subunit, a second *Nde*I target was generated at the translational initiation site of the *sucD* gene so that the *sucD* gene could be independently subcloned into the expression vector using a 2.2 kb *Nde*I–*Bam*HI fragment. The 'two-primer' method was used with Φ1N as the template and S48 as the mutagenic primer (see the Materials and methods section). This primer binds to positions 5806–5782 in the *sucC–sucD* junction except for the triple mismatch that generates on ATA → CAT mutation immediately upstream of the *sucD* coding region (Buck *et al.*, 1985). One of several progeny phages containing the desired sequence was designated Φ1NN. In addition to changing the sequence between the ribosome binding site and initiator codon of *sucD*, the *sucC* gene is extended by four codons into the *sucD* coding region (Fig. 3). Both changes could affect the expression and activity of the corresponding SCS. This has not been investigated and it is possible that the use of Φ1NN may be limited to mutation and independent overexpression of the α subunit, rather than as a universal template for the mutagenesis and overexpression of both subunits, either independently or in combination. Nevertheless, Φ1N and Φ1NN are now available as versatile mutagenic templates providing cassettes for the overexpression of the α and β subunits (*Nde*I_1–*Bam*HI), the β subunit (*Nde*I_1–*Sal*I) and the α subunit (*Nde*I_2–*Bam*HI), in pJLA503.

Site-directed mutagenesis of SCS

Several *sucC* and *D* mutations were created in preliminary studies aimed at testing the importance of some potential active site residues. The active site of SCS is thought to be formed in the region of α–β contact; the ATP-binding site and the histidine residue (His-246α) that is phosphorylated being located in the α subunit, whereas the binding sites for succinate and CoA reside in the β subunit (Nishimura, 1986). There is evidence from chemical modification studies that a second histidine residue is important for stabilizing the reactive thiol form of a cysteine residue in the CoA binding site of the β subunit (Collier & Nishimura, 1979). It has also been found that the enzyme is inactivated by modification of one of the three tryptophan residues in the β subunit (Ybarra *et al.*, 1986). Inspection of the primary structure revealed several putative CoA binding sites (Buck *et al.*, 1985), and it may be significant that two of these are associated with residues of the putative active site types: Trp-43β, Cys-47β and His-50β.

Mutagenic oligonucleotides were used with Φ1 or Φ1N to direct the following amino acid substitutions in the α or β subunits of SCS encoded by the corresponding M13 derivatives: His-246α(CAC) → Asp(GAC) in Φ1-M3 and Φ1N-M3; Trp-43β(TGG) → Val(GTG) in Φ1N-M4; Cys-47β(TGT) → Ser(TCT) in Φ1N-M5; and His-50β(CAC) → Ala(GCA) in Φ1N-M6 (see the Materials and methods section). The substitutions were chosen to minimize steric effects and the His → Asp change was introduced with the ultimate aim of being able to effect the separation of mixtures of wild-type, mutant and hybrid enzymes, according to their charge. In each case, between two and six out of six potential mutants had the desired sequence, and representatives were tested for their ability to express IPTG-inducible SCS activity in extracts of infected cultures of *E. coli* JM101, and for significant amplification of the α and β subunits in the SDS/PAGE profiles of the crude extracts (Table 3). As expected, no increase in SCS activity was observed following infection with Φ1-M3, which incorporates asparate rather than histidine at the phosphorylation site; indeed it appeared to lower the basal level expressed by the host. However, the failure to detect the inducible amplification of both SCS subunits with Φ1-M3 was unexpected. A plasmid (pGS201) in which the mutant 0.5 kb *Sal*I fragment was used to replace the correspond-

Table 3. Effects of site-directed mutation of the *sucC* and *D* genes on the SCS activities expressed by M13 derivatives

Cultures of *E. coli* JM101 were infected with derivatives of M13mp10 containing wild-type mutant *sucC* and *D* genes and their SCS activities were measured in the supernatant fraction of ultrasonic extracts after growth for 16 h in the presence or absence of the *lac* inducer IPTG (200 μg/ml). The extracts were analysed by SDS/PAGE for inducible amplification (+) of polypeptides corresponding to the α and β subunits. Specific activities are expressed as μmol of substrate transformed/h per mg of protein.

Source of *sucC* and *D* genes	Amino acid substitution	SCS specific activity (μmol/h per mg of protein)		Amplification of α and β subunits
		Uninduced	Induced	
M13mp10	Vector	10	11	–
Φ1	Wild-type	15	42	++
Φ1-M3	His-246α → Asp	6	6	–
Φ1N-M4	Trp-43β → Val	9	9	++
Φ1N-M5	Cys-47β → Ser	9	25	++
Φ1N-M6	His-50β → Ala	9	10	+

ing fragment of pGS131 and was then resequenced, also failed to express detectable SCS activity. It is not known why amplification of the SCS subunits was not detected with Φ1-M3, though several reasons could be advanced. For example, it is possible that the mutation affects the folding of the α subunit, making it more susceptible to proteolytic degradation, or that correct folding (or even phosphorylation) of the α subunit may be essential for assembly of the $\alpha_2\beta_2$ tetramer. In either case, failure to form the tetrameric enzyme may lead to subunit degradation or to the production of sedimenting aggregates.

The three phages with missense mutations in the *sucC* gene exhibited amplified expression of both subunits, Φ1N-M6 being somewhat weaker than the others (Table 3). However, there was no significant amplification of SCS activity with Φ1N-M4 and Φ1N-M6 which is consistent with Trp-43 and His-50 being essential residues of the β subunit. In contrast, there was an increase with Φ1N-M5, albeit less than the wild-type control, and this indicates that Cys-47 is not an essential residue of the β subunit. Clearly further studies are needed with mutant enzymes that have been overexpressed in a host which is devoid of wild-type enzyme. The *kdp-suc* deletion strain, TK3D18 (Buck *et al.*, 1986), is not entirely satisfactory because a promoterless $sucC^+D^+$ region is retained and thus available for recombination with the mutant genes, and because the strain grows very poorly due to the absence of several other citric acid cycle enzymes. A useful strain could be constructed by replacing the *sucABCD* operon of a wild-type strain with a *sucAB* version that had been deleted *in vitro* for the *sucCD* region.

Nevertheless, the results illustrate the potential value of the strategy for mutagenesis and thus far, they support predictions that Trp-43β and His-50β may be essential residues at the CoA binding site. Cys-47β appears to be less important and another cysteine residue in the same region of the primary structure (Cys-25β) may be functional.

Nucleotide sequence of an unidentified gene adjacent to the *suc* operon

Previous maxicell studies with plasmids containing segments of the *suc* region pointed to the existence of an unidentified gene immediately downstream of the *suc* operon (Buck *et al.*, 1986). The unassigned gene had the opposite polarity to the *suc* operon and was designated *g30* because it encodes a polypeptide of M_r 30000 (hereinafter designated P30). The nucleotide sequence of *g30* was deduced from M13 clones containing fragments of the *SalI–BglII* (Sa_2–Bg_5) segment of the *suc* region

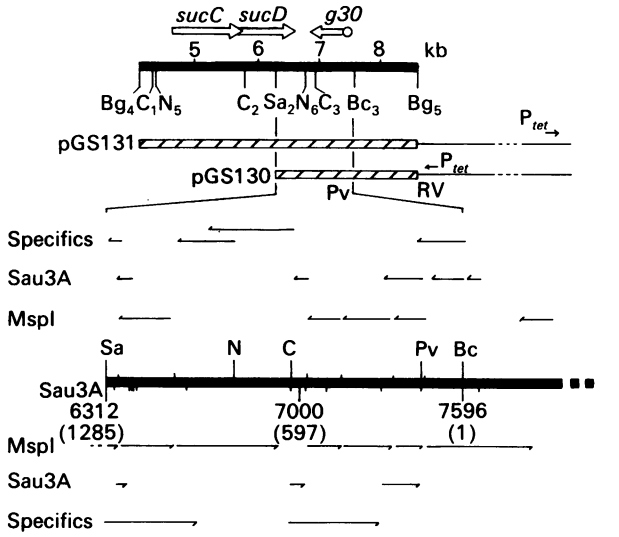

Fig. 5. Location of the unidentified gene (*g30*) and summary of DNA sequence data from M13 clones

A partial restriction map of the *sucCD* region indicating the position and polarity of *g30* and the segments of bacterial DNA (hatched bars) cloned into the *tet* gene of pBR322 are shown. The polarities of the *tet* promoters are indicated by the horizontal arrows and the *PvuI* (Pv) and *Eco*RV (RV) sites of pGS130 referred to in the text are included. Other restriction sites for *BclI*(Bc), *BglII*(Bg), *ClaI*(C), *NruI*(N) and *SalI*(SA) are abbreviated and numbered according to Buck *et al.* (1985). The expanded region shows the positions and extents of sequence data obtained from M13 clones. The nucleotide positions are numbered from an arbitrary *TaqI* site upstream of the *suc* operon and the corresponding positions in the complementary strand (see Fig. 6) are given in parentheses starting at the *BclI* site (Bc_3). The coding region of *g30* extends from positions 6773 to 7492 (105–824).

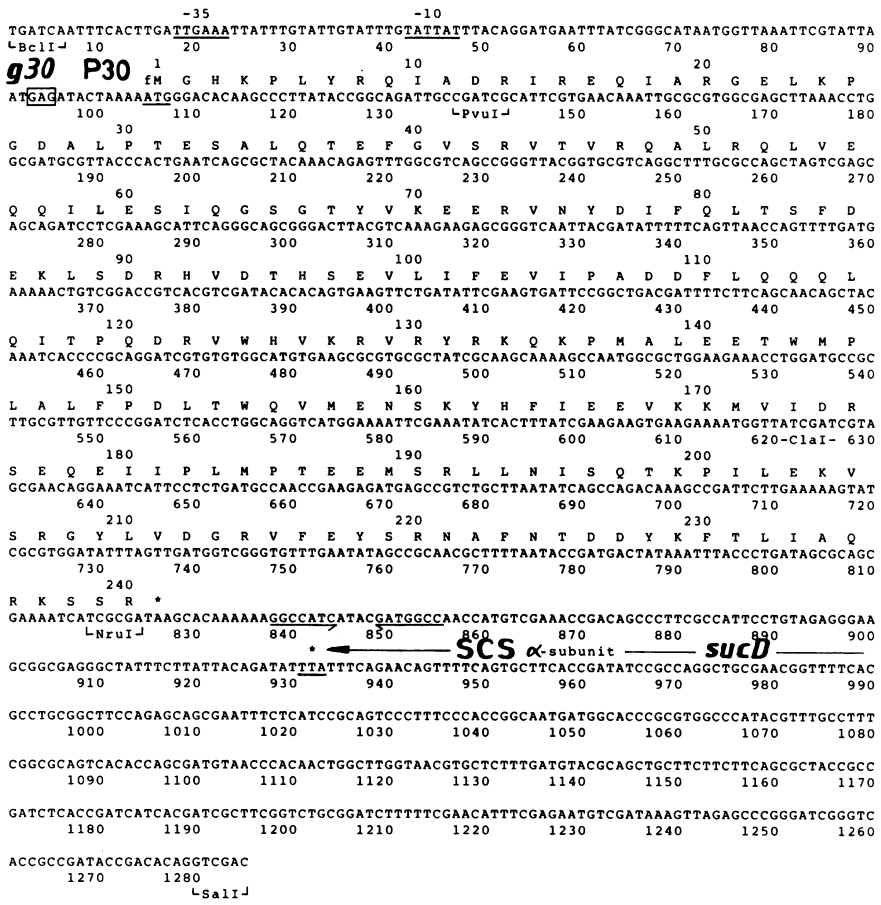

Fig. 6. Nucleotide sequence of the unidentified *g30* gene and the primary structure of its putative product P30

The nucleotide positions are numbered from the first base of the *BclI* site to the last base of the *SalI* site (Bc_3–Sa_2 in Fig. 5; the corresponding positions in the complementary strand are 7596 and 6312). The amino acid sequence is numbered from the initiating formylmethionine. Potential ribosome binding site (boxed) and promoter sequences (underlined, −35 and −10) are indicated. The region of hyphenated dyad symmetry that might be a transcriptional terminator for *g30* and the *suc* operon is underscored by converging arrows and translational stop sites for the two genes are marked by asterisks. Relevant restriction sites are also indicated.

(Fig. 1). The direction and extent of sequencing for individual clones is shown in Fig. 5 and the nucleotide sequence of the 1285 bp *SalI*–*BclI* segment (Sa_2–Bc_3) is presented with a *BclI* to *SalI* polarity in Fig. 6. The sequence is fully overlapping and derived from both DNA strands. It overlaps and extends the previously published sequence of 13061 bp by 818 bp to 13879 bp. A further 875 bp of preliminary sequence data for the adjacent 1050 bp *BclI*–*BglII* (Bc_3–Bg_5) segment was also obtained. Apart from the 3′ end of *sucD*, only one potential coding region was detected with the FRAMESCAN program of Staden & McLachlan (1982). This coding region converges on the *suc* operon, it starts with an AUG codon at position 105, ends with a UAA codon at position 825, and encodes a polypeptide of M_r 28271 (Fig. 6). The location, polarity and size of gene product are all in good agreement with those predicted for *g30*.

Features of the nucleotide sequence. The *g30* start codon is located in a region corresponding to a translational initiation site when analysed by the PERCEPTRON algorithm of Stormo *et al.* (1982). A computer search for promoter sequences revealed a potential promoter at positions 19–48, which is identical to the *E. coli* consensus at ten out of twelve positions (Fig. 6). The sequence was also examined for regions of hyphenated dyad symmetry, and the most significant is located at position 838–855 [$\Delta G = -72$ kJ/mol (-17.2 kcal/mol); Tinoco *et al.*, 1973]. This sequence was previously identified as a potential *rho*-independent terminator of the *suc* operon, and it is conceivable that it could serve as transcriptional terminator for both the *suc* operon and *g30* (Fig. 6).

The codon usage for *g30* (Table 4) shows that a fairly high proportion (6%) of modulatory codons and a relatively low proportion (40%) of optimal energy codons in the diagnostic set are used (Grosjean & Fiers, 1982). This pattern of codon usage is typical of weakly-expressed genes encoding low-abundance proteins such as repressors, so that *g30* product might be a regulatory protein.

Properties of the putative *g30* gene product (P30). The primary structure of the protein (P30) encoded by *g30* is shown in Fig. 6. It contains 230 amino acids and is of M_r 27251 (including the initiating formylmethionine). The amino acid composition is typical of *E. coli* proteins

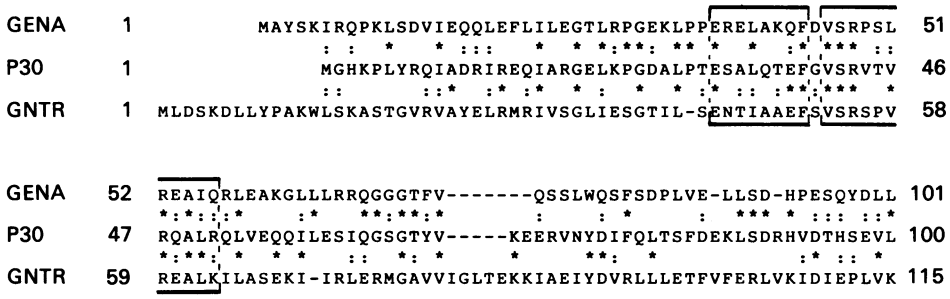

Fig. 7. Alignment of the amino acid sequences of three gene products with unknown or uncertain function

The *N*-terminal segments of P30 and GENA, the products of the unidentified genes (*g30* and *geneA*) that are adjacent to the *sucABCD* and *aceEF-lpd* operons of *E. coli* (respectively), and GNTR, a putative repressor of the gluconate operon (*gntRKPZ*) of *B. subtilis*, are aligned for maximum similarity. Asterisks signify identical residues and colons denote conserved substitutions scoring $\geqslant 0.1$ in the MDM_{78} mutation data matrix (Schwartz & Dayhoff, 1978). The position of the potential λCro-like DNA-binding structure (helix-turn-helix) is indicated.

Table 4. Codon usage in the *g30* gene

The codon pairs enclosed in boxes are those whose use varies significantly between strongly and weakly expressed genes, and asterisks denote potential modulatory codons (Grosjean & Fiers, 1982). The initiating formylmethionine codon is included.

UUU	F	8	UCU	S	0	UAU	Y	5	UGU	C	0
UUC	F	2	UCC	S	0	UAC	Y	3	UGC	C	0
UUA	L	4	UCA	S	2	UAA	-	0	UGA	-	0
UUG	L	2	UCG	S	4	UAG	-	0	UGG	W	3
CUU	L	5	CCU	P	2	CAU	H	1	CGU	R	8
CUC	L	2	CCC	P	2	CAC	H	4	CGC	R	6
*CUA	L	3	CCA	P	2	CAA	Q	5	*CGA	R	2
CUG	L	6	CCG	P	5	CAG	Q	14	*CGG	R	4
AUU	I	8	ACU	T	2	AAU	N	4	AGU	S	2
AUC	I	6	ACC	T	7	AAC	N	1	AGC	S	7
*AUA	I	2	ACA	T	3	AAA	K	8	*AGA	R	0
AUG	M	7	ACG	T	1	AAG	K	6	*AGG	R	0
GUU	V	4	GCU	A	3	GAU	D	11	GGU	G	1
GUC	V	6	GCC	A	1	GAC	D	3	GGC	G	4
GUA	V	1	GCA	A	0	GAA	E	17	*GGA	G	2
GUG	V	7	GCG	A	6	GAG	E	5	*GGG	G	1

except for a relatively low abundance of serine, alanine and glycine, and a higher glutamate plus glutamine content than normal. The hydropathy profile is characteristic of soluble globular proteins (results not shown).

Attempts to assign a function to *g30* and its product involved unsuccessful searches for similar sequences in several nucleic acid and protein databanks using the EMBLSCAN and PEPSCAN programs (Bishop & Thompson, 1984; Bishop, 1984). However, a Dap Protein Database Search (Collins *et al.*, 1988), kindly performed by Dr. J. F. Collins, was more successful. The data base contained over 8000 protein sequences and highly significant similarities were detected between the *N*-terminal regions of P30 and two other proteins of unknown or uncertain function, GENA and GNTR. For P30 and GENA it was estimated that 1.5×10^6 databases of the same size would have to be analysed before such an alignment would be expected to occur by chance, and for P30 and GNTR, 2.5×10^3 databases would have to be searched. GENA is the product of an unidentified but non-essential gene (*geneA*) located upstream of the *aceEF-lpd* operon that encodes the pyruvate dehydrogenase complex of *E. coli* (Stephens *et al.*, 1983), and GNTR is a putative transcriptional repressor encoded by the *gntR* gene in the gluconate operon (*gntRKPZ*) of *Bacillus subtilis* (Fujita *et al.*, 1986). The *g30* product (M_r 27251, 220 residues) resembles the products of *geneA* (M_r 27049, 236 residues) and *gntR* (M_r 28277, 243 residues) in size, and the two *E. coli* genes have codon usages that are typical of weakly-expressed genes. Sequence alignments for the *N*-terminal segments are shown in Fig. 7. The similarity between P30 and GENA (30% identical, 55% similar when conservative substitutions are included) is slightly greater and more sustained than for P30 and GNTR (28%, 51%) over the region considered, and GENA and GNTR are the least similar (18%, 41%).

If the proposed repressor function of GNTR is confirmed, the observed sequence similarities suggest that the three proteins belong to the same family of regulatory proteins. Indeed, the similar regions may represent structurally analogous DNA-binding domains. A search was made for potential helix-turn-helix structures resembling those found in many DNA-binding proteins, and one region, residues 31–51 in P30, had some of the features of the λCro-like structures (Fig. 7). Using the weight matrix of Dodd & Eagan (1987), the scores for the relevant regions of GENA, P30 and GNTR correspond to respective probabilities of 60%, 7% and < 1% that they are λCro-like. This analysis is therefore consistent with GENA and P30 being potential regulatory proteins.

Conclusion

A successful strategy for site-directed mutagenesis of the SCS of *E. coli* using M13 derivatives, and for its thermoinducible overexpression from λ promoters, has been developed. This should be invaluable for studying many of the interesting questions concerning the assembly and catalytic activity of this enzyme, and more especially because the solution of its three-dimensional structure is imminent. In the course of this study, insights into the need for a closely co-ordinated expression of both subunits have emerged, and enzyme produc-

tion would appear to require a correctly-folded or phosphorylated α subunit. Further studies on the separate overexpression of individual subunits and with mutants having different amino acid substituents at the site of phosphorylation are clearly needed to resolve these possibilities. The preliminary mutagenic studies have also supported predictions about the identities of histidine and tryptophan (but not cysteine) residues that may be located at the CoA-binding site in the β subunit.

A potentially interesting finding emerging from the nucleotide sequence of the unidentified gene (*g30*) that is adjacent to the *sucABCD* operon, is the partial sequence similarity between its gene product (P30), and a putative repressor of the *B. subtilis* gluconate operon (*gntRKPZ*), and more especially with the product (GENA) of an unidentified gene (*geneA*) that is adjacent to the *aceEF-lpd* operon. The *aceEF-lpd* operon is located at 2.8 min in the *E. coli* linkage map and it encodes the specific E1p and E2p components of the pyruvate dehydrogenase complex, and also the E3 component that is common to pyruvate and 2-oxoglutarate dehydrogenase complexes. The *sucABCD* operon at 16.8 min encodes the specific E1o and E2o components of the 2-oxoglutarate dehydrogenase complex and the β and α subunits of SCS. The transcriptional relationships so far established for the two operons are as follows:

geneA-aceEF-lpd, and *sucAB-sucCD-g30*.

The transcripts of the adjacent genes and the possibility of readthrough from *geneA* to the *ace* operon has not been investigated. Furthermore, there is no experimental evidence linking *geneA* and *g30* with the regulation of the adjacent operons. However, there are several reasons for considering this to be feasible, e.g. the clustering of functionally-related genes, the common transcriptional terminator that has been proposed for the *suc* operon and *g30*, and the need for regulatory mechanisms to control the expression of the two multienzyme complexes and to co-cordinate the expression of the *lpd* gene with the synthesis of the two complexes. The results clearly identify the need for a detailed functional characterization of the two potential regulatory genes.

This work was supported by the Science and Engineering Research Council. We are indebted to: Dr. M. E. Spencer, Dr. J. S. Miles, Dr. S. Spiro and Dr. S. C. Andrews for helpful discussions; and G. D. Searle for synthetic oligonucleotides; Dr. J. C. Collins for kindly performing the protein database search; and Dr. J. E. G. McCarthy for providing the expression vector.

REFERENCES

Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 3963–3965
Bishop, M. (1984) Bioessays **1**, 29–31
Bishop, M. & Thompson, E. (1984) Nucleic Acids Res. **12**, 5417–5474
Boyer, H. W. & Roulland-Dussoix, D. J. (1969) J. Mol. Biol. **41**, 459–472
Bridger, W. A. (1974) Enzymes (3rd Edn.) **10**, 581–606
Bridger, W. A., Ramaley, R. G. & Boyer, P. D. (1969) Methods Enzymol. **13**, 70–75
Buck, D., Spencer, M. E. & Guest, J. R. (1985) Biochemistry **24**, 6245–6252
Buck, D., Spencer, M. E. & Guest, J. R. (1986) J. Gen. Microbiol. **132**, 51–77
Collier, G. E. & Nishimura, J. S. (1979) J. Biol. Chem. **254**, 10925–10930
Collins, J. C., Coulson, A. F. W. & Lyall, A. (1988) Comput. Appl. Biosci. **4**, 67–71
Coulson, A. F. W., Collins, J. C. & Lyall, A. (1987) Computer J. **30**, 420–424
Darlison, M. G., Spencer, M. E. & Guest, J. R. (1984) Eur. J. Biochem. **141**, 351–359
Dodd, I. B. & Eagan, J. B. (1987) J. Mol. Biol. **194**, 557–564
Fujita, Y., Fujita, T., Miwa, Y., Nihashi, J. & Aratani, Y. (1986) J. Biol. Chem. **261**, 13744–13753
Grosjean, H. & Fiers, W. (1982) Gene **18**, 199–209
Henning, W. D., Upton, C., McFadden, G., Majumdar, R. & Bridger, W. A. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 1432–1436
Kunkel, T. A. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 488–492
Kramer, B., Kramer, W. & Fritz, H.-J. (1984) Cell (Cambridge, Mass.) **38**, 879–887
Laemmli, U. K. (1970) Nature (London) **227**, 680–685
Lennox, E. S. (1955) Virology **1**, 190–206
Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) J. Biol. Chem. **193**, 265–275
Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
Messing, J. (1979) Recomb. DNA Techn. Bull. **2**, 43–48
Messing, J. & Vieira, J. (1982) Gene **19**, 269–276
Miles, J. S. & Guest, J. R. (1985) Nucleic Acids Res. **13**, 131–140
Miles, J. S. & Guest, J. R. (1987) Biochem. Soc. Symp. **54**, 45–65
Miller, J. H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, New York
Nishimura, J. S. (1986) Adv. Enzymol. **58**, 141–172
Norrander, J., Kempe, T. & Messing, J. (1983) Gene **26**, 101–106
Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980) J. Mol. Biol. **143**, 161–178
Schauder, B., Blöcker, H., Frank, R. & McCarthy, J. E. G. (1987) Gene **52**, 279–283
Schwartz, R. M. & Dayhoff, M. D. (1978) in Atlas of Protein Sequences and Structure (Dayhoff, M. O., ed.), vol. 5, suppl. 3, pp. 353–358, National Biomedical Research Foundation, Washington
Spencer, M. E. & Guest, J. R. (1985) Mol. Gen. Genet. **200**, 145–154
Spencer, M. E., Darlison, M. G., Stephens, P. E., Duckenfield, I. K. & Guest, J. R. (1984) Eur. J. Biochem. **141**, 361–374
Staden, R. (1979) Nucleic Acids Res. **6**, 2601–2611
Staden, R. (1980) Nucleic Acids Res. **8**, 3673–3694
Staden, R. & McLachlan, A. D. (1982) Nucleic Acids Res. **10**, 141–156
Stephens, P. S., Darlison, M. G., Lewis, H. M. & Guest, J. R. (1983) Eur. J. Biochem. **133**, 155–162
Stormo, G. D., Schneider, T. D., Gold, L. & Ehrenfeucht, A. (1982) Nucleic Acids Res. **10**, 141–156
Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Ullenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) Nature (London) New Biol. **246**, 40–41
Weitzman, P. D. J. (1981) Adv. Microb. Physiol. **22**, 185–244
Wolodko, W. T., Brownie, E. R. & Bridger, W. A. (1980) J. Bacteriol. **143**, 231–237

Wolodko, W. T., James, M. N. G. & Bridger, W. A. (1984) J. Biol. Chem. **259**, 5316–5320
Wolodko, W. T., Kay, C. M. & Bridger, W. A. (1986) Biochemistry **25**, 6245–6252
Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) Gene **33**, 103–119
Ybarra, J., Prasad, A. R. S. & Nishimura, J. S. (1986) Biochemistry **25**, 7174–7178
Zoller, M. J. & Smith, M. (1983) Methods Enzymol. **100**, 468–500
Zoller, M. J. & Smith, M. (1984) DNA **3**, 479–488

Received 9 November 1988/27 January 1989; accepted 13 February 1989