Molecular cloning, sequencing and expression of cytochrome c_2 from *Rhodospirillum rubrum*

Susan J. SELF,* C. Neil HUNTER† and Robin J. LEATHERBARROW*‡

*Imperial College of Science, Technology and Medicine, South Kensington, London SW7 2AY, and †Department of Biochemistry, University of Sheffield, Sheffield S10 2TN, U.K.

Cytochrome c_2 (M_r 12840) of the purple photosynthetic bacterium *Rhodospirillum rubrum* functions as a mobile electron carrier in the cyclic photosynthetic electron-transport system of this organism. It acts as the electron donor to photochemically oxidized reaction centres and is reduced in turn by electrons from the cytochrome bc_1 complex. By using synthetic oligonucleotides based on the known amino acid sequence of the protein, the structural gene (*cycA*) has been identified and isolated. DNA sequence analysis indicates the presence of a typical prokaryotic 23-residue signal sequence, suggesting that the protein is synthesized as a precursor which is processed during its secretion into the periplasm. Evidence is presented for the production of assembled cytochrome c_2 in *Escherichia coli*, but recombinants grow poorly and are unstable, suggesting toxicity of the gene product in this organism.

INTRODUCTION

Cytochrome c_2 forms a part of the cyclic photosynthetic electron-transport system of purple bacteria such as *Rhodospirillum* (*Rs.*) *rubrum* and is located on the periplasmic aspect of the intracytoplasmic membranes that have the photosynthetic apparatus. It has been shown (Daldal *et al.*, 1986) that cytochrome c_2 is not essential for photosynthetic growth of *Rhodobacter* (*Rb.*) *capsulatus*, whereas it is essential in photoheterotrophic growth of *Rb. sphaeroides* (Donohue *et al.*, 1988).

The soluble cytochrome, c_2 , has 112 amino acid residues (M_r 12840) and contains a protohaem IX group covalently linked to the polypeptide backbone by thioether bridges at the conserved site Cys-Xaa-Yaa-Cys-His. The primary amino acid sequence has been determined (Dus *et al.*, 1968) and shows substantial sequence similarity to the primary structures of mammalian cytochromes c (Dus *et al.*, 1968; Dickerson & Timkovich, 1975). The main difference arises from a short insertion in c_2 at residue 53. Despite this difference, the crystal structure of cytochrome c_2 , solved to 0.2 nm (2 Å) resolution (Salemme *et al.*, 1973a), shows a striking resemblance to that of eukaryotic cytochrome c (Salemme *et al.*, 1973b). This conservation in primary and tertiary structures has led to the suggestion that aerobic respiration in prokaryotic and eukaryotic organisms has evolved from a common ancestor. Unlike eukaryotic cytochromes c , *Rs. rubrum* cytochrome c_2 does not undergo an extensive conformational change upon oxidation–reduction. Overall it is most similar in structure to ferrocycytochrome c .

We report here the use of synthetic oligonucleotides to identify the structural gene for *Rs. rubrum* cytochrome c_2 , *cycA*. The DNA sequence of the cloned gene is presented and is compared with the sequences of the *cycA* genes from *Rb. sphaeroides* (Donohue *et al.*, 1986)

and *Rb. capsulatus* (Daldal *et al.*, 1986). The gene product is shown to be expressed in *Escherichia coli*. Constructs expressing cytochrome c_2 are found to grow very poorly and are unstable, suggesting that the gene product is toxic *in vivo*.

MATERIALS AND METHODS

Strains and plasmids

E. coli strain TG2 (Gibson, 1984), a *recA*[−] derivative of TG1 [K12, $\Delta(lac-pro)$, *supE*, *thi*, *hsdD5/F'* *traD36*, *proA*⁺*B*⁺, *lacI*⁺, *lacZ* Δ M15], was used for all cloning procedures. *Rs. rubrum* strain S1 was grown and used for DNA isolation. Plasmids pUC9 (Vieira & Messing, 1982) pUC119 (Vieira & Messing, 1987) and pACYC184 (Chang & Cohen, 1978) were employed in the cloning steps. Antibiotics were used at final concentrations of 50 µg/ml (chloramphenicol) and 15 µg/ml (tetracycline) where appropriate. Bacteriophage M13mp19 or pUC plasmids and derivatives were propagated on L-broth plates additionally supplemented with 1.3 mM-isopropyl β -D-thiogalactoside and 5-bromo-4-chloroindol-3-yl β -D-galactoside (270 µg/ml).

Materials

Restriction enzymes were kindly donated by GeneSys (Feltham, Middx., U.K.), except for *Hind*III, which was bought from Pharmacia. Northumbria Biologicals (Cramlington, Northd., U.K.) supplied the polynucleotide kinase used to end-label oligonucleotide probes. Amersham International were suppliers of all radiochemicals, and also of Hybond-N nylon filters used for Southern blots and colony/plaque screening. Sequenase was from Cambridge Bioscience, and GeneClean from Stratech Scientific, Luton, Beds., U.K. Random hexamers (pdN₆) used in oligolabelling were from Pharmacia.

‡ To whom correspondence and reprint requests should be sent.

These sequence data will appear in the EMBL/GenBank/DBJ Nucleotide Sequence Databases.

DNA analysis and cloning strategy

Rs. rubrum chromosomal DNA was prepared from a 1-litre culture of cells grown photoheterotrophically in Sistrom's (1960) minimal medium A. Cells were harvested at 5000 rev./min in a Sorvall SS-34 rotor, washed in 10 mM-Tris/HCl/1 mM-EDTA, pH 8, and resuspended in the same buffer containing lysozyme at 1 mg/ml. After 15 min on ice, proteinase K was added to 60 µg/ml and the lysed cells were incubated at 50 °C for 16 h. CsCl was added to the extent of 1 g/ml of solution, together with ethidium bromide to a final concentration of 0.1 mg/ml. DNA was banded at 48000 rev./min in a Sorvall Ti70 rotor over a period of at least 16 h. Banded chromosomal DNA was then collected and purified as described previously (Maniatis *et al.*, 1982).

Two synthetic oligonucleotides, P1 and P2 (Fig. 1), each with built-in redundancies in sequence, were used to identify and isolate the gene. The oligonucleotides were synthesized on an Applied Biosystems model 380B DNA synthesizer and purified by using a Pharmacia f.p.l.c. system, elution from the MonoQ column being achieved with a linear gradient of 0–0.5 M-NaCl in 10 mM-NaOH. When used to probe Southern blots (see below), P1 and P2 were 5'-end-labelled with [γ -³²P]ATP. The reaction mixture contained 100 pmol of oligonucleotide in 50 mM-Tris/HCl (pH 8.0)/10 mM-MgCl₂/5 mM-dithiothreitol/0.1 mM-EDTA and 0.2 unit of polynucleotide kinase in a 50 µl reaction volume. Before use, radiolabelled probes were purified through a spinning column of Sephadex G-25 in TNES [0.14 M-NaCl/0.02 M-Tris/HCl (pH 7.6)/0.1 % SDS] (Maniatis *et al.*, 1982).

DNA digested with restriction enzymes was resolved by electrophoresis through agarose gels using either Tris/borate/EDTA (TBE) or Tris/acetate/EDTA (TAE) as a buffer system (Maniatis *et al.*, 1982). Restriction fragments were recovered either by electroelution from agarose/TBE gels (Maniatis *et al.*, 1982) or from agarose/TAE using GeneClean according to the manufacturer's directions.

DNA fragments resolved by gel electrophoresis were blotted on to Amersham Hybond-N in 20 × SSC (3 M-NaCl/0.3 M-sodium citrate) after acid depurination, alkali denaturation and neutralization under the conditions recommended by the manufacturer. Similarly, plaques or colonies transferred to Hybond-N filters were denatured and neutralized *in situ*. Blots were prehybridized at 65 °C

in the buffer system described by Amersham International. Radiolabelled probes were added and hybridization was carried out at 65 °C for 10 min, followed by cooling to room temperature over several hours. Non-specifically bound probe was washed off with 6 × SSC at a temperature of 53 °C for 10 min, and the blots were exposed to preflashed X-ray film at –70 °C using an intensifying screen.

The *cycA* gene from *Rb. sphaeroides* has been cloned (C. N. Hunter, unpublished work) as part of a DNA fragment which restores the ability to photosynthesize to a *Rb. sphaeroides* cytochrome *c*₂ deletion mutant (Donohue *et al.*, 1986). The *Rb. sphaeroides cycA* gene was used to generate probes for *Rs. rubrum cycA*. A restriction fragment containing the gene was radioactively labelled with [α -³²P]dATP by oligonucleotide labelling (Feinberg & Vogelstein, 1983).

Single-stranded M13 DNA was sequenced by the dideoxy-chain-termination method using Sequenase under conditions described by the manufacturer. Two sets of reactions were allowed to proceed: dGTP-containing mixes were used initially, and dITP was substituted for dGTP to resolve areas of compression in the sequence. Preliminary DNA sequence data enabled the design of the two sequencing primers, SQ1 and SQ2, used in later sequencing reactions.

Cytochrome *c*₂ was purified from *Rs. rubrum* essentially as described previously (Bartsch, 1971).

RESULTS AND DISCUSSION

Cloning *cycA* from *Rs. rubrum*

Two oligonucleotide probes (P1 and P2) were designed to hybridize to the coding strand of the *cycA* gene, based on the known amino acid sequence of cytochrome *c*₂ (Dus *et al.*, 1968). Both oligonucleotides were designed as 'best guesses' according to the codon usage of *Rs. rubrum* based on the published DNA sequence of ribulose-1,5-bisphosphate carboxylase/oxygenase (Nargang *et al.*, 1984). Each oligonucleotide had sequence redundancy built in at three positions. Oligonucleotide P1 was a 29-mer, designed to hybridize to the region encoding Glu-54–Thr-63, whereas oligonucleotide P2, 35 bases in length, was based on the amino acid sequence near the C-terminus from Met-91 to Glu-102 (Fig. 1).

The isolated genomic DNA was digested with a series

Fig. 1. Amino acid sequence and position of hybridization of the oligonucleotides P1 and P2, used to identify and isolate the *cycA* gene from *Rs. rubrum*

The protein sequence is that published by Dus *et al.* (1968).

Fig. 2. Southern analysis of genomic DNA from *Rs. rubrum*

Various restriction digests were probed with ^{32}P -labelled oligonucleotides P1 and P2 (Fig. 1). The filters were washed at 53 °C in 6×SSC to remove non-specifically bound label, and the X-ray film was exposed at -70 °C to reveal the bands: a, DNA digested with *Bam*HI; b, DNA digested with *Bam*HI and *Hind*III; c, DNA digested with *Bam*HI and *Pst*I; d, undigested DNA. The positions of *Cl*aI-digested λ -phage size markers are not shown.

of restriction endonucleases, and the fragments were separated on a 1% (w/v)-agarose gel in TBE, together with DNA from λ -phage digested with *Hind*III as size markers. Both oligonucleotides were found to hybridize to a single *Hind*III fragment of approx. 3.5 kb. In addition, a double digest using *Hind*III and *Bam*HI resulted in a fragment of approx. 0.5 kb, to which P1 hybridized strongly, and a fragment of approx. 2 kb, to which P2 hybridized (Fig. 2). This shows that a *Bam*HI site is present in the central portion of the gene, which was valuable to the cloning strategy employed subsequently.

A *Hind*III digestion of chromosomal *Rs. rubrum* DNA was resolved on an agarose gel, and fragments between 3 and 4 kb were recovered. Ligation of these restriction fragments into plasmid pUC9 was followed by transformation into the competent *E. coli* strain TG2, so generating a mini-library of *Rs. rubrum* DNA. The resulting recombinant colonies were screened with the

Fig. 3. Restriction map of 3.3 kb *Hind*III fragment containing the *cycA* gene

radiolabelled P1 probe, and positively hybridizing colonies were replated. However, these colonies grew very slowly, rapidly lost viability and appeared to lyse on the plate. This is possibly indicative of a toxic gene product. After several days, reddish-brown discoloration of the cells and surrounding medium could be observed. However, owing to the unstable nature of the isolates, which would not grow in liquid culture, these constructs were not suitable for further study.

By using the same *Hind*III digest, a second mini-library was generated in pACYC184 (Chang & Cohen, 1978) and screened with radiolabelled oligonucleotide P1 as described above. It was hoped that the absence of any strong promoter sequences in pACYC184 would result in a more stable construct. Colonies hybridizing to oligonucleotide P1 were also probed with radiolabelled P2. A single recombinant colony was selected (containing plasmid pSS1) which hybridized strongly to both P1 and P2. This plasmid additionally hybridized with restriction fragments containing the *cycA* gene from *Rb. sphaeroides*, confirming that the *Rs. rubrum cycA* gene was present in the pSS1 construct. Single and double restriction digests of pSS1 were resolved by gel electrophoresis, enabling

Fig. 4. Sequencing the *cycA* gene

The region of DNA sequenced is indicated as a line, with the coding region emphasized. The arrows represent individual gel readings, and key restriction sites are shown (B, *Bam*HI; Hc, *Hinc*II).

Fig. 6. Second-derivative visible-region spectra of (a) an extract from cells containing the *cycA* gene cloned into pUC119 and (b) purified cytochrome c_2 from *Rs rubrum*

Arrows mark the positions of peak maxima at 416, 521 and 550 nm in both samples. *E. coli* TG2 cells resuspended to the same absorbance did not show maxima at these positions (results not shown).

as a *Bam*HI fragment of approx. 6 kb. By using universal primer, and sequencing primers SQ1 and SQ2, approx. 650 bp was sequenced. Data were obtained for both coding and non-coding strands, as illustrated in Fig. 4. The consensus nucleotide sequence is shown in Fig. 5. The predicted sequence is in good agreement with the published amino acid sequence, except at positions 45 and 73, which are both aspartic acid residues, and not asparagine residues as previously described. In addition there is a 23-amino-acid sequence between the first in-frame ATG codon and the known start of the gene. This sequence is composed of hydrophobic and charged residues, characteristics typical of prokaryotic signal sequences. Similar precursor sequences have also been found in the *cycA* genes of *Rb. sphaeroides* and *Rb. capsulata* (Daldal *et al.*, 1986; Donohue *et al.*, 1986). The DNA sequence shows 55% sequence similarity to that of *Rb. sphaeroides* and 53% to that of *Rb. capsulata*, whereas the amino acid sequences are 43 and 44% similar respectively, a reflection of the common ancestry of these photosynthetic organisms. Upstream of the initiation codon, at -8 to -12 bp, is the presumed ribosome-binding site, GAGGA. The gene ends with a

single termination codon, with a potential hairpin loop, T1, centred 55 bp downstream of the TAG termination codon. This region is G + C-rich, capable of forming a loop with a ΔG of $-62.8 \text{ kJ } (-15 \text{ kcal}) \cdot \text{mol}^{-1}$ (Tinoco *et al.*, 1973), and has a series of thymine residues immediately 3' to it. These characteristics are typical of the factor-independent terminators described by Platt (1986). An *E. coli* consensus promoter sequence (Hawley & McClure, 1983) is found 170–130 bp upstream of the initiating methionine residue (Fig. 5). Such a sequence has also been found in the ATPase operon of *Rs. rubrum* (Falk & Walker, 1985) and more recently as the second of two promoter sequences controlling transcription of the *puh* gene (Bérard *et al.*, 1989). This second promoter ($P_{\text{puh}2}$) was found to direct transcription in *E. coli* and is thought to control low-level constitutive expression of the gene, with a strong oxygen-regulated promoter $P_{\text{puh}1}$ upstream of it. Donohue *et al.* (1986) found the *cycA* gene in *Rb. sphaeroides* produces two mRNA transcripts. By analogy, it therefore seems likely that the *E. coli* consensus promoter sequence found in *Rs. rubrum cycA* directs constitutive expression of the gene, and it is possible that a strong oxygen-regulated promoter lies upstream of the sequence region.

Expression in *E. coli*

The *cycA* gene was isolated on a 1.1 kb *Sma*I fragment and subcloned into pUC119 (Vieira & Messing, 1987), the transformed cells being plated out on L-broth plates supplemented with ampicillin. In this way we attempted to verify our initial observations that *cycA* cloned into a plasmid containing a promoter would express cytochrome c_2 protein. After several days, several colonies were observed to become reddish-brown in colour, as had been found previously. These colonies could not be grown in liquid culture, and only grew very poorly on solid media. Once again, the growth was too poor to allow isolation of sufficient DNA for analysis. Colonies were recovered from a plate and resuspended in 0.1 M-sodium phosphate buffer, pH 6.95, two or three crystals of sodium dithionite were added, and spectra recorded on a Shimadzu UV/3000 spectrophotometer. The second-derivative spectrum (Fig. 6) shows absorption maxima at 416, 521 and 550 nm, compared with the published values for cytochrome c_2 of 415, 521 and 550 nm (Horio & Kamen, 1961). The second-derivative spectrum of cytochrome c_2 isolated from *Rs. rubrum* is shown for comparison. Spectra of *E. coli* TG2 resuspended to the same absorbance and recorded under the same conditions failed to show any such maxima. The recombinant protein also showed the changes in absorbance on oxidation which are characteristic of cytochrome c_2 , although lysozyme treatment of the cells was necessary before the fully oxidized spectrum could be recorded. These spectra indicate expression of cytochrome c_2 in *E. coli* and suggest that the discoloration observed is due to the accumulation of the cytochrome c_2 gene product. Stable expression of the gene will enable further experimentation to show the degree of processing of the gene product in *E. coli* and the mode of attachment of the haem moiety to the polypeptide backbone.

We thank Jack Knill-Jones for synthesizing the oligonucleotides used in this work, and Dr. Duncan Clark of GeneSys Ltd. for donating the restriction enzymes. Thanks also go to Dr. Tim Wells of Smith, Kline and French Research

Ltd., The Frythe, Welwyn, Herts., U.K., for help with the spectroscopy. We are grateful to Dr. Timothy Donohue for the use of his *Rb. sphaeroides cycA* deletion mutant.

REFERENCES

- Bartsch, R. G. (1971) *Methods Enzymol.* **23**, 344–363
- Bérard, J., Bélanger, G. & Gingras, G. (1989) *J. Biol. Chem.* **262**, 10897–10903
- Chang, A. C. Y. & Cohen, S. C. (1978) *J. Bacteriol.* **134**, 1141–1156
- Daldal, F., Cheng, S., Applebaum, J., Davidson, E. & Prince, R. C. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2012–2016
- Dickerson, R. E. & Timkovich, R. (1975) *Enzymes* 3rd Ed. **9A**, 397–547
- Donohue, T. J., McEwan, A. G. & Kaplan, S. (1986) *J. Bacteriol.* **168**, 962–972
- Donohue, T. J., McEwan, A. G., Van Doren, S., Crofts, A. R. & Kaplan, S. (1988) *Biochemistry* **27**, 1918–1925
- Dus, K., Sletten, K. & Kamen, M. D. (1968) *J. Biol. Chem.* **243**, 5507–5518
- Falk, G. & Walker, J. E. (1985) *Biochem. J.* **229**, 663–668
- Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13
- Gibson, T. J. (1984) Ph.D. Thesis, University of Cambridge
- Hanahan, D. (1985) in *DNA Cloning* (Glover, D. M., ed.), vol. 1, pp. 109–135, IRL Press, Oxford
- Hawley, W. R. & McClure, D. K. (1983) *Nucleic Acids Res.* **11**, 2237–2255
- Horio, T. & Kamen, M. D. (1961) *Biochim. Biophys. Acta* **48**, 266–286
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Nargang, F., McIntosh, L. & Somerville, C. (1984) *Mol. Gen. Genet.* **193**, 220–224
- Platt, T. (1986) *Annu. Rev. Biochem.* **55**, 339–372
- Salemme, F. R., Freer, S. T., Xuong, Ng. H., Alden, R. A. & Kraut, J. (1973a) *J. Biol. Chem.* **248**, 3910–3921
- Salemme, F. R., Kraut, J. & Kamen, M. D. (1973b) *J. Biol. Chem.* **248**, 7701–7716
- Sistrom, W. R. (1960) *J. Gen. Microbiol.* **22**, 778–785
- Tinoco, I., Jr., Borer, P. N., Dengler, B. & Levine, M. D. (1973) *Nature (London) New Biol.* **246**, 40–41
- Vieira, J. & Messing, J. (1982) *Gene* **19**, 259–268
- Vieira, J. & Messing, J. (1987) *Methods Enzymol.* **153**, 3–11

Received 15 June 1989/6 September 1989; accepted 4 October 1989