Biochimica et Biophysica Acta, 1143 (1993) 109–111

BBABIO 40270

Short Sequence-Paper

PCR cloning, sequence analysis and expression of the *cyb*C genes encoding soluble cytochrome *b*-562 from *Escherichia coli* B strain OP7 and K strain NM522

Michael K. Trower

MRC Molecular Genetics Unit, MRC Centre, Cambridge (UK)

(Received 1 March 1993)

Key words: Cytochrome *b*-562; Electron transport protein; Heme protein; Periplasmic protein; *cyb*C gene; PCR; (*E. coli*)

The *cyb*C genes encoding cytochrome *b*-562 from B and K strains of *Escherichia coli* were shown to differ in size following their isolation by in vitro amplification using the polymerase chain reaction. Nucleotide sequencing of the genes and flanking regions revealed the discrepancy was due to a 26 bp deletion at the 5′ end of the K strain sequence that included the initiation codon. Expression studies confirmed that the K strain gene was untranslated. These data indicate that the *cyb*C gene product is non-essential in *E. coli*.

Cytochrome *b*-562 from *Escherichia coli* is a small (M_r 12000), haem-containing, soluble protein [1]. The protein mediates electron transport, though its functional role and associated redox partners have yet to be identified. Analysis of its deduced primary structure, determined from a B strain *cyb*C DNA sequence, revealed a typical periplasmic leader sequence of 22 amino acids which is cleaved to generate a 106 amino acid mature protein [2]. The structure of the mature protein has been highly characterised by X-ray crystallography and forms a left twisted 4-α-helical bundle [3].

For structure/function purposes I sought to isolate the *E. coli cyb*C gene encoding cytochrome *b*-562. A primer pair, whose design was based upon the reported DNA sequence, was used to amplify by PCR the homologous regions from 3 genomic DNA samples prepared from *E. coli* B strain OP7 (Sydney Brenner culture collection, Cambridge, UK) and the K strains NM522 and DH5α. Intriguingly, the size of the amplified fragments from the K strains was approximately 20–30 bp smaller than the anticipated product size of 650 bp generated using the *E. coli* B strain DNA as template. To identify the location and nature of this discrepancy the products from both the B strain OP7 and NM522 K strain amplifications were subcloned and then sequenced. Comparisons between these sequences and the previously reported B strain *cyb*C DNA sequence are shown in Fig. 1. Remarkably, the reduced size of the PCR product generated from the K strain genomic DNA is due to 26 bp deletion, which begins almost immediately after the putative ribosomal binding site and continues into the periplasmic signal sequence, thereby including the start codon. No other possible initiation codon appears to be present to enable translation of the structural gene. However, in spite of this deletion the deduced coding sequence of the remaining K strain *cyb*C gene is very conserved when compared to the B-strain sequences. There are only 11 base changes, 3 of which lead to amino acid substitutions. Interestingly, these alterations in sequence are clustered in two 'hot' spots: the first is broad and is located around the signal peptide cleavage processing site; five of six base alterations are in the 3rd position and only one generates a conserved amino acid change of Ile to Val. The other is close to the C-terminus and is highly focused consisting of 5 base changes in an 11 base sequence with two further conserved amino acid alterations, an Ala to Ser and Gln to Lys.

The regions that regulate transcription of the *cyb*C gene have not been clearly identified, though a weak promotor lies 5′ to a *Sma*I site (marked on Fig. 1) in the B strain sequence [2]. Since the initially amplified

Correspondence to: M.K. Trower, Biochemical Targets Dept., Microbiology Division, Glaxo Group Research Ltd., Greenford, Middlesex UB6 0HE, UK.

EMBL accession numbers for the *E. coli cyb*C genes encoding cytochrome *b*-562 are X67290 for the OP7 B strain and S67289 for the NM522 K strain.

Abbreviation: PCR, polymerase chain reaction.

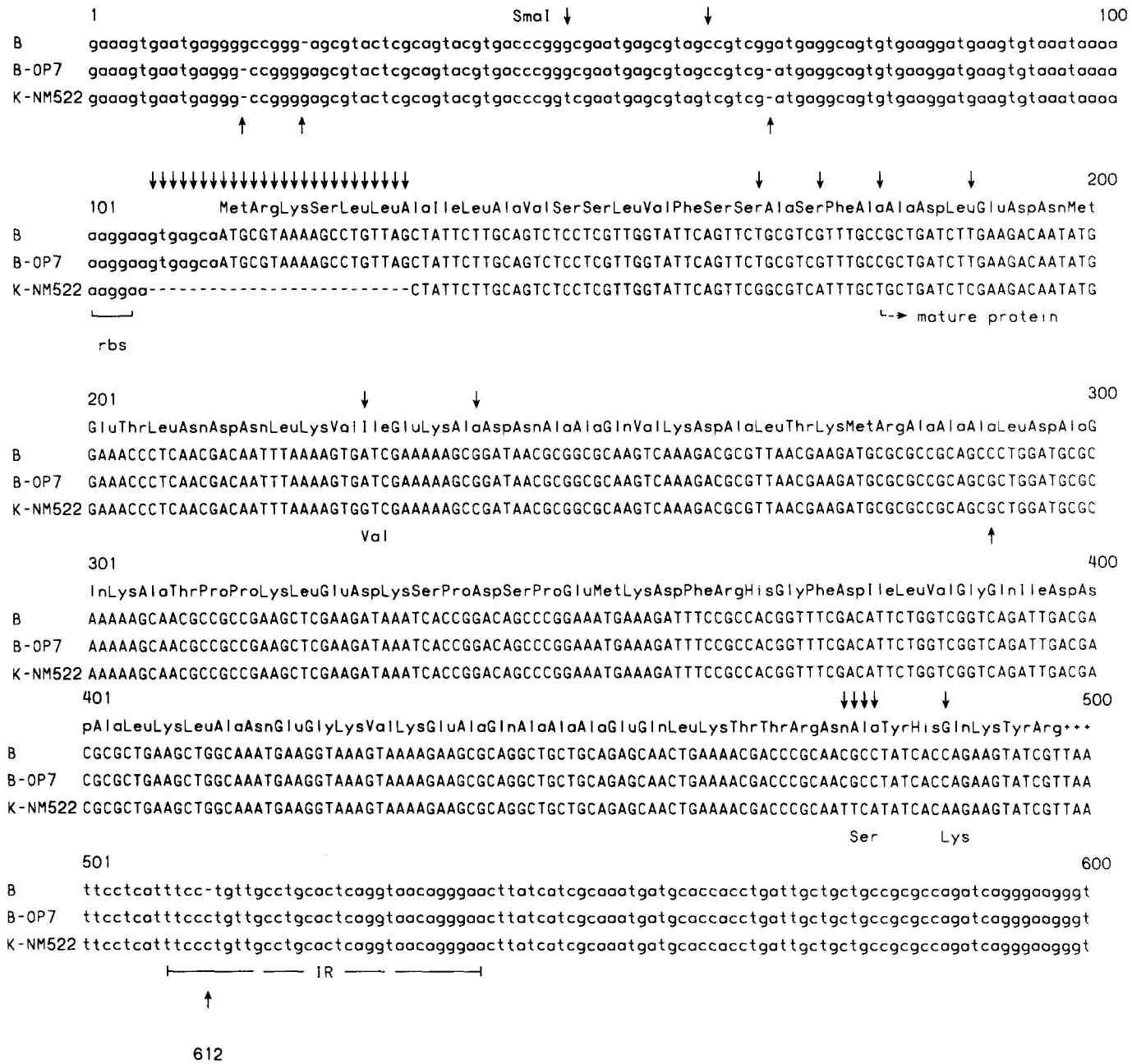

Fig. 1. Alignments between *cyb*C genes and flanking DNA sequences from *E. coli* B and K strains. Sequences are as reported by Ref. 2 (B), and herein for the B strain OP7 (B-OP7) and K strain NM522 (K-NM522). The DNA segments for B-OP7 and K-NM522 were prepared with genomic DNA's as templates for PCR between the forward primer, 237F; 5′ CGC ACT GCA TCG AGA CGG 3′ and reverse primer, 238R 5′ CCC CGT CGA CGG CAA ATT TGT G (contains a *Sal*I site at the 5′ end flanked by 4 C's) using the proofreading thermostable pfu DNA polymerase (Stratagene). Reactions, carried out in duplicate for each genomic DNA template, in a total volume of 50 μl contained: forward primer (10 μM), 5 μl; reverse primer (10 μM), 5 μl; pfu reaction buffer×10 (as supplied by manufacturer), 5 μl; dNTP's (2 mM of each base), 5 μl; genomic DNA (15–20 μg/ml), 1 μl; water, 28.5 μl and pfu DNA polymerase (2.5 U/μl), 0.5 μl. The reactions were heat denatured for 5 min before they were 'hot started' by the addition of the pfu enzyme. The following program was then immediately initiated for a total of 28 cycles: denaturation, 95°C for 0.5 min; primer annealing, 50°C for 0.5 min; primer extension, 72°C for 2 min. After the final cycle an additional step of 72°C for 5 min was added to ensure full length double stranded products. Products generated from independent PCR reactions for each genomic DNA template, were blunt-end cloned into *Eco*RV cut pBluescript II KS(+) and sequenced. The deduced primary structure for the B-strain *cyb*C genes are shown above the DNA sequences; amino acids that differ in the K strain are shown below. Differences between the B strain and K strain nucleotide sequences are shown by a downward arrow (↓) above the sequences, alterations between the B strain nucleotide sequence reported by [2] and the two sequences described herein are shown by an upward arrow (↑) below the sequences. Putative DNA regulatory regions, the ribosomal binding site (rbs) and a 3′ prime imperfect inverted repeat (IR) that may form a RNA polymerase termination signal are underlined. The signal peptide cleavage site is also indicated on the figure.

segments contain these regulatory regions a further primer set were used to amplify by PCR and clone the B and K strain structural genes to attempt their expression under control of the powerful *lac* promoter.

Conditions are the same as described in the legend for Fig. 1; the oligonucleotides 243F, CCC GGG CGA ATG AGC GTA and 238R, 5′ CCC CGT CGA CGG CAA ATT TGT G were used a PCR primers, and the in vitro amplified products were again blunt-end ligated into *Eco*RV cut pBluescript II KS(+). The orientation of the amplified segments with respect to the *lac* promotor were determined by restriction mapping. Cytochrome *b*-562 contents of cell free extracts prepared from overnight cultures induced with 1 mM IPTG, were determined spectrophotometrically [1]. The haem protein was only detectable in extracts from cells harbouring plasmids carrying the B strain *cyb*C gene in the correct orientation with respect to the promoter (3.2% of total protein). Cytochrome *b*-562 could not be detected in cells carrying the K strain gene placed in either orientation, providing further evidence that the observed deletion prevents its translation.

These results suggest that cytochrome *b*-562 encoded by the *cyb*C gene is a non-essential protein in *E. coli*, though it is possible that its functional role in the K-strain examined may have been taken over by another protein within the periplasmic space. These data also exemplify a potential problem that may be encountered when using laboratory strains of *E. coli* for gene isolation from either prepared genomic libraries or by the use of PCR. Since laboratory strains of this species have been obtained by mutagenic methods, the described deletion could have been generated during preparation of a common K strain parent. However, it is also feasible that the deletion is a feature of K strains of *E. coli*; the completion of the *E. coli* genome project currently being undertaken at the University of Wisconsin, Madison, WI, USA should clarify the situation.

I would like to thank Dr. Sydney Brenner for his help, encouragement and suggestions and E.I. duPont de Nemours and Co. Inc., Wilmington, DE, USA for financial support. The genomic NM522 DNA was a kind gift from Greg Elgar in the Molecular Genetics Unit.

References

1 Itagaki, E. and Lowell, P.H. (1966) J. Biol. Chem. 241, 3687–3695.
2 Nikkila, H., Gennis, R.B. and Sligar, S.G. (1991) Eur. J. Biochem. 202, 309–313.
3 Lederer, F. Glatigny, A., Bethge, P.H., Bellamy, H.D. and Mathews, F.S. (1981) J. Mol. Biol. 148, 427–448.