Volume 212, number 1, 168–172 FEB 04432 February 1987

The azurin gene from *Pseudomonas aeruginosa* codes for a pre-protein with a signal peptide

Cloning and sequencing of the azurin gene

G.W. Canters

Gorlaeus Laboratories, State University Leiden, PO Box 9502, 2300 RA Leiden, The Netherlands

Received 1 December 1986

The azurin gene from *Pseudomonas aeruginosa* is located on a 1.3 kb long *Pst*I DNA fragment. Its nucleotide sequence has been determined. It appears that the gene codes for a pre-protein with a 19 amino acid long signal sequence which possibly assists in the transport of the azurin over the periplasmic membrane.

Azurin; Nucleotide sequence; Signal peptide; Pre-protein; (*Pseudomonas aeruginosa*)

1. INTRODUCTION

Despite considerable theoretical and experimental effort, much is still unknown about the molecular details of biological electron transfer. The question of how the coupling between redox partners takes place and is regulated has been addressed with partial success in only a few cases [1–5]. Whether electron transfer by a particular redox protein is vectorial and whether electron transport proteins possess separate entries and exits for the electrons they are supposed to transport are still largely matters of speculation and debate [6]. Likewise, little is known about the 'pathway' of electrons inside redox proteins; even the applicability of such a concept may be questionable.

The increase over the past 10 years in the number of redox proteins of which the 3-dimensional structure has been reported, however, now makes these questions gradually more amenable to experimental research. The cytochromes occupy a prominent position in this respect among the redox proteins and more recently attention has also been focussed on the relatively small (10–20 kDa) blue copper proteins. Structures of plastocyanin under various conditions have been reported [7], as well as the 3-dimensional structures of the azurins from *Pseudomonas aeruginosa* [8] and *Alcaligenes denitrificans* [9]. Determination of the structures of various other blue copper proteins (amicyanins and pseudo-azurins) is under way [10].

An attractive strategy to study the mode of operation of redox proteins is to try to observe the effect of small structural variations on their function. Since rec DNA techniques have been shown to be particularly powerful in this respect [11], research aimed at site-directed mutagenesis of blue copper proteins was started in this laboratory. Attention has been focussed thereby on the azurin from *P. aeruginosa*. This choice was motivated by the fact that this is by far the most extensively studied prokaryotic blue copper protein to date. Its 3-dimensional structure was reported in 1978 [8] and its redox kinetics have been studied exhaustively [12]. The hypothesis has been formulated that 2 areas on the protein surface may be involved in electron transfer, namely a

Correspondence address: G.W. Canters, Gorlaeus Laboratories, State University Leiden, PO Box 9502, 2300 RA Leiden, The Netherlands

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

hydrophobic patch around copper ligand His-117 and a second region close to the more buried His-35 [12,13]. The latter residue has also been implicated in the pH dependence of the spatial conformation of the protein and in the pH dependence of its redox activity, as observed in the electron transfer reaction with cytochrome *c*-551 [14–17]. Furthermore, the relation between the mode of coordination of the redox center (the Cu atom) and the redox properties of the protein has intrigued many researchers [18]. There is, therefore, a rich choice of residues that appear suitable and relevant candidates for site-directed mutagenesis.

As a first step in this direction the azurin gene must be isolated. Here we report on the successful cloning and sequencing of the azurin gene from *P. aeruginosa*. The method chosen to identify the gene consists of the hybridization of bacterial DNA with synthetic, radiolabelled oligonucleotide probes [19]. During the sequencing of the DNA it became clear that the gene codes for a pre-azurin with a 19 amino acid long signal sequence. The implication of this finding is briefly discussed.

2. MATERIALS AND METHODS

2.1. *DNA isolation*

Chromosomal DNA was isolated from *P. aeruginosa* strain CIT135, which was a kind gift from Dr A.P.F. Turner, Cranfield Institute of Technology, England. Bacteria were grown anaerobically as described [20]. DNA was isolated from 3 g of (wet) bacterial paste according to [21]. Repeated treatment with pronase and prolonged phenol/$CHCl_3$ extraction proved necessary to obtain a purity index (A_{260}/A_{280}) of better than 1.7.

2.2. *Labelling*

Labelling of oligonucleotide probes was performed by phosphorylation of about 25 pmol oligonucleotide with about 30 pmol [γ-^{32}P]ATP (spec. act. 3000 Ci/mmol) according to [22].

For gel electrophoresis of DNA 1–2% agarose minigels were used. Subsequent screening of dried [19] and denatured [19] gels with radioactive probes was performed by preincubation for at least 1 h at 45°C with 1 × Denhardt's solution [22], also containing 0.1 mM ATP, 1 mM Na_2HPO_4, 1 mM $Na_2P_2O_7$, 0.8 M NaCl, 200 μg/ml of partially hydrolysed RNA, 0.1 M Tris (pH 7), 6 mM EDTA and 5 ml NP40 per l. After preincubation the radioactive probe was added and incubation was continued for at least 3 h more during which the temperature was allowed to drop to 20–25°C. Following hybridization, gels were washed 3 times at room temperature with 6 × SSC (saline sodium citrate buffer [22]) and incubated for 10 min in 6 × SSC, 0.1% SDS and 10 min in 6 × SSC, both times at 41°C (probe 3) or 35°C (probes 1,2). Autoradiograms were obtained after 2–72 h of exposure at −70°C with the use of 1 or 2 intensifying screens.

2.3. *Plasmid isolation*

Isolation of plasmids was performed according to Birnboim and Doly [23]. Restriction digestions and ligations were performed as in [22]. Transformation of *E. coli* KMBL1164 with pBR322 and *E. coli* JM101 with M13 was completed as described [22]. DNA sequencing was performed according to the dideoxy method of Sanger et al. [24].

3. RESULTS

The oligonucleotide probes used for the screening experiments are shown in fig.1. The probes were kindly synthesized by Professor J.H. van Boom and co-workers and consisted of mixtures of 8 tetradecamers (probes 1,2) or 16 heptadecamers (probe 3). The hybridization procedure was optimized on restriction digests of chromosomal DNA. The best results were obtained with *Pst*I and *Xho*I digests as shown in fig.2. Two bands at around 1.3 and 2.3 kb in the *Pst*I digest and at around 7.1 and 4.1 kb in the *Xho*I lane were observed in the autoradiogram (fig.2). Hybridization with probes 1 and 2 appeared to be less selective, but resulted in positive signals in the autoradiogram at the 1.3 kb position in the *Pst*I lane. Further work was concentrated, therefore, on this 1.3 kb *Pst*I fragment.

A band around 1.3 kb was cut from an agarose gel and the extracted DNA cloned into the *Pst*I site of plasmid pBR322. After transformation into *E. coli* and plating on LB agar, Tet^r and Amp^s clones were selected, cultured individually on LB medium and processed in groups of 12 clones. Plasmid isolations of each group were screened with probe 3 for insertion of the azurin gene. In this way 144 clones could be screened in one run. After screen-

aa sequence	probe
- Met - Glu - Phe - Asn - Thr - (13 … 17)	1. 5' d-GT PTT PAA YCT CAT
- Met - Phe - Phe - Cys - Thr - (109 … 113)	2. 5' d-GT PCA PAA PAA CAT
- Met - Gly - His - Asn - Trp - Val - (44 … 49)	3. 5' d-AC CAA PTT PTG NCC CAT

Fig.1. Oligonucleotide probes 1, 2 and 3 (top to bottom) as used in the hybridization experiments (right). The corresponding regions of the amino acid sequence are shown at the left. P, purine; Y, pyrimidine nucleotide; N, all 4 nucleotides.

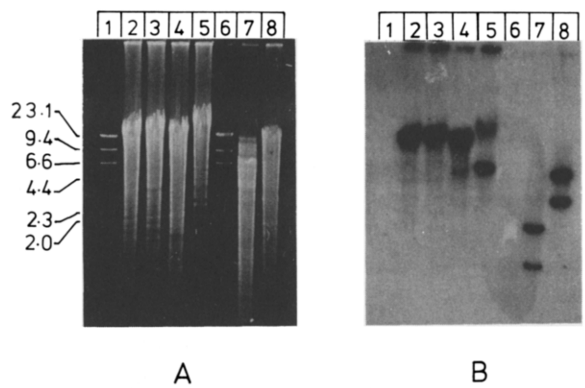

Fig.2. Gel electrophoresis (1% agarose) of digests of *P. aeruginosa* chromosomal DNA (A). 1–2 μg DNA was applied per lane. Lanes: 1, DNA marker; 2, *Bam*HI; 3, *Bgl*II; 4, *Eco*RI; 5, *Hin*dIII; 6, DNA marker; 7, *Pst*I; 8, *Xho*I. Autoradiogram of the same gel after hybridization with probe 3 is shown on the right (B). Conditions: hybridization at 20°C o.n., wash at 41°C in 6 × SSC, 0.1% SDS for 20 min and subsequently at 41°C in 6 × SSC for 20 min; 21 h exposure at −70°C with 2 intensifying screens.

ing of about 300 clones, a positive group of 12 clones was found, and by repeating the screening procedure for the individual clones a single positive clone was identified rapidly.

The inserted fragment appeared to produce strong hybridization signals not only with probe 3, but with the other 2 probes as well, indicating that a large portion, if not the entire gene, was located on the pBR322 inserted *Pst*I fragment.

This fragment was subsequently screened for restriction sites by incubation with various restriction enzymes. At this stage again effective use could be made of probes 1–3 for rapid screening of the digests and location of the restriction sites with respect to the azurin gene. Digestion with *Acc*I produced 2 *Pst*I/*Acc*I fragments of about 650 and 150 bp and 1 *Acc*I/*Acc*I fragment of about 500 bp. The two larger fragments were cloned into M13mp19 and used for subsequent sequencing experiments. Next an *Hae*III digest of the 1.3 kb *Pst*I

```
                     -20                                    -13
                     Met - Leu - Arg - Lys - Leu - Ala - Ala - Val -
TGCCTAGGAGGCTGCTCC   ATG   CTA   CGT   AAA   CTC   GCT   GCG   GTA   -37

-12
Ser - Leu - Leu - Ser - Leu - Leu - Ser - Ala - Pro - Leu - Leu - Ala -
TCC   CTG   CTG   TCC   CTG   CTC   AGT   GCG   CCG   CTG   CTG   GCT   -1

1
Ala - Glu - Cys - Ser - Val - Asp - Ile - Gln - Gly - Asn - Asp - Gln -
GCC   GAG   TGC   TCG   GTG   GAC   ATC   CAG   GGT   AAC   GAC   CAG   36

13
Met - Gln - Phe - Asn - Thr - Asn - Ala - Ile - Thr - Val - Asp - Lys -
ATG   CAG   TTC   AAC   ACC   AAT   GCC   ATC   ACC   GTC   GAC   AAG   72

25
Ser - Cys - Lys - Gln - Phe - Thr - Val - Asn - Leu - Ser - His - Pro -
AGC   TGC   AAG   CAG   TTC   ACC   GTC   AAC   CTG   TCC   CAC   CCC   108

37
Gly - Asn - Leu - Pro - Lys - Asn - Val - Met - Gly - His - Asn - Trp -
GGC   AAC   CTG   CCG   AAG   AAC   GTC   ATG   GGC   CAC   AAC   TGG   144

49
Val - Leu - Ser - Thr - Ala - Ala - Asp - Met - Gln - Gly - Val - Val -
GTA   CTG   AGC   ACC   GCC   GCC   GAC   ATG   CAG   GGC   GTG   GTC   180

61
Thr - Asp - Gly - Met - Ala - Ser - Gly - Leu - Asp - Lys - Asp - Tyr -
ACC   GAC   GGC   ATG   GCT   TCC   GGC   CTG   GAC   AAG   GAT   TAC   216

73
Leu - Lys - Pro - Asp - Asp - Ser - Arg - Val - Ile - Ala - His - Thr -
CTG   AAG   CCC   GAC   GAC   AGC   CGC   GTC   ATC   GCC   CAC   ACC   252

85
Lys - Leu - Ile - Gly - Ser - Gly - Glu - Lys - Asp - Ser - Val - Thr -
AAG   CTG   ATC   GGC   TCG   GGC   GAG   AAG   GAC   TCG   GTG   ACC   288

97
Phe - Asp - Val - Ser - Lys - Leu - Lys - Glu - Gly - Glu - Gln - Tyr -
TTC   GAC   GTC   TCC   AAG   CTG   AAG   GAA   GGC   GAG   CAG   TAC   324

109
Met - Phe - Phe - Cys - Thr - Phe - Pro - Gly - His - Ser - Ala - Leu -
ATG   TTC   TTC   TGC   ACC   TTC   CCG   GGC   CAC   TCC   GCG   CTG   360

121
Met   Lys   Gly   Thr   Leu   Thr   Leu   Lys
ATG   AAG   GGC   ACC   CTC   ACC   CTG   AAG   TGA   TGC   GCG   ACA   396
```

Fig.3. Nucleotide sequence of the *P. aeruginosa* azurin gene together with the amino acid sequence [27]. The region between black bars (nucleotide −73/−68) is the purported Shine-Dalgarno sequence. Restriction sites are underlined as follows: singly, *Acc*I/*Taq*I; doubly, *Hae*III; triply, *Taq*I.

fragment was cloned into mp10 and two clones which appeared positive with probe 2 were used for further sequencing. Finally, the 650 bp *Acc*I/*Pst*I fragment was subfractionated with *Taq*I and cloned directly into mp10. Clones with a 220 and a 300 bp insert were used for further sequencing. The final results of the sequencing experiments are shown in fig.3.

4. DISCUSSION

An unexpected finding in the present study was the occurrence of two strong signals in the autoradiogram when probe 3 was applied in the hybridization experiments with restriction digests of chromosomal DNA (see fig.2). It is conceivable that the bacterial chromosome codes for more than one blue copper protein. The induction of different blue copper proteins in one organism through the application of different media and growth conditions has been reported for a number of methylotrophic bacteria, for instance [25,26]. Experiments of this kind have as yet not been described for *P. aeruginosa*, however.

The nucleotide sequence in fig.3 appears in agreement with the known amino acid sequence [27] (apart from the signal peptide, of course). The restriction sites inferred from the digestion and hybridization experiments are confirmed by the results in fig.3, i.e. a single *Acc*I site at amino acids 22/23, *Taq*I sites at amino acid positions 22/23 and 97/98 and an *Hae*III site at amino acid positions 45/46. Additional *Hae*III sites appear at positions 67/68 and 116/117.

While the end of the gene is clearly signalled by a stop codon, it came as a surprise that a start codon was lacking at the anticipated position directly in front of the first amino acid of the azurin (Ala). Instead an in-frame AUG triplet 60 nucleotides upstream from this position was found. It is assigned as the start codon on the basis of two observations: a 7 nucleotide long Shine-Dalgarno consensus sequence (underlined in fig.3) occurs at the expected position (7 nucleotides upstream from the purported start codon). Secondly, the 19 amino acids upstream of the mature azurin, as deduced from the nucleotide sequence, exhibit all the characteristics of a prokaryotic signal peptide [28]. Preliminary sequencing results on the region further upstream of the start codon (up to 60 nucleotides) did not reveal any other in-frame or out-of-frame start codon. (No transcription start sequences ('−10' and '−35 boxes') could be located either.)

Until now no signal peptides have been reported for the blue copper proteins, except for plastocyanin [29], which is a eukaryotic electron transporter, however. Its signal sequence is accordingly much more complex, since it must program for the passage over more than one membrane. The length of the presently found signal sequence (19 amino acids) concurs with the average length of 20 amino acids observed for prokaryotic signal peptides [28]. The required presence of 1 or (more often) 2 positively charged residues at the N-terminus is confirmed (Arg and Lys at positions −18 and −19), as well as the presence of a β-turn-promoting residue (Pro or Gly) in the first half of the peptide (Pro at −4) and the presence of a serine at position −6 [30,31]. The small residues usually observed at the processing site are also seen here (Ala, Ala at positions −1, +1) as well as a roughly 10 amino acids long central hydrophobic core (positions −7 to −17 in the present case). An interesting observation is the occurrence of 2 serines in this hydrophobic region. It is often assumed that this part of the peptide may adopt an α-helical conformation, at least when the signal peptide is embedded in a membrane [32]. The presence of 2 serines seems partly to counteract this tendency.

Finally, the present results may shed some light on the functional role of the azurin. The presence of a signal sequence in front of the azurin almost certainly means that the protein is transported over the plasma membrane into the periplasmic space. The occurrence of blue copper proteins in this cellular compartment in other organisms has been demonstrated recently for the amicyanins in some methylotrophic bacteria [25,33]. Whether periplasmic targeting is a general characteristic of the azurins cannot be stated from the present results. The most that can be said at the moment is that the occurrence of an Ala residue at amino acid position 1 in all the azurins sequenced up till now [27] is compatible with a processing site at that position [28,30–32].

In view of their relatively high redox potentials bacterial blue copper proteins are usually placed towards the end of bacterial electron transport

chains. Circumstantial evidence has been reported that the azurin from *P. aeruginosa* acts as the electron donor of the terminal nitrite oxidase [34]. A location of this enzyme in the outer membrane or on the periplasmic side of the plasma membrane would make the necessity for the azurin to be present in the periplasmic space understandable.

ACKNOWLEDGEMENTS

The author thanks Professor L. Bosch for his continuous encouragement and advice, and thanks him and Dr C.W.A. Pleij for providing him with experimental facilities in the Biochemical Laboratory. Frequent experimental advice and help from Dr K. Rietveld, Dr E. Vijgenboom and Dr J. Brouwer is gratefully acknowledged. The author thanks Professor J.H. van Boom and his co-workers for the synthesis of the oligonucleotide probes, Dr H.L. Heineker (San Francisco) for experimental advice during the early stages of this work, and Professor J. Reedijk for critically reading the manuscript.

REFERENCES

[1] Bosshard, R.B., Davidson, R.W., Knaff, D.B. and Millett, F. (1986) J. Biol. Chem. 261, 190–193.
[2] Hazzard, J.T. and Tollin, G. (1985) Biochem. Biophys. Res. Commun. 130, 1281–1286.
[3] Guerlesquin, F., Noailly, M. and Bruschi, M. (1985) Biochem. Biophys. Res. Commun. 130, 1102–1108.
[4] Stonehuerner, J., O'Brien, P., Geren, L., Millett, F., Steidl, J., Yu, L. and Yu, C.-A. (1985) J. Biol. Chem. 260, 5392–5398.
[5] King, G.C., Binstead, R.A. and Wright, P.E. (1985) Biochim. Biophys. Acta 806, 262–271.
[6] Bechtold, R., Kuehn, C., Lepre, C. and Isied, S.S. (1986) Nature 322, 286–288.
[7] Church, W.B., Guss, J.M., Potter, J.J. and Freeman, H.C. (1986) J. Biol. Chem. 261, 234–237.
[8] Adman, E.T. and Jensen, L.H. (1981) Isr. J. Chem. 21, 8–12.
[9] Norris, E.N., Anderson, B.F. and Baker, E.N. (1986) J. Am. Chem. Soc. 108, 2784–2785.
[10] Hormel, S., Adman, E.T., Walsh, K.A., Beppu, T. and Titani, K. (1986) FEBS Lett. 197, 301–304.
[11] Fersht, A.R., Leatherbarrow, R.J. and Wells, N.C.T. (1986) Trends Biochem. Sci. 11, 321–325.
[12] Adman, E.T. (1985) in: Metalloproteins (Harrison, P.M. ed.) ch.1, pp.1–41, Verlag Chemie, Weinheim.
[13] Groeneveld, C.M. and Canters, G.W. (1985) Eur. J. Biochem. 153, 559–564.
[14] Silvestrini, M.C., Brunori, M., Wilson, M.T. and Darley-Usmar, V.M. (1981) J. Bioinorg. Chem. 14, 327–338.
[15] Adman, E.T., Canters, G.W., Hill, H.A.O. and Kitchen, N.K. (1982) FEBS Lett. 143, 287–292.
[16] Corin, A.F., Bersohn, R. and Cole, P.E. (1983) Biochemistry 22, 2032–2038.
[17] Ugurbil, K., Norton, R.S., Allerhand, A. and Bersohn, R. (1977) Biochemistry 16, 886–894.
[18] Gray, H.B. and Malmström, B.G. (1983) Comments Inorg. Chem. 2, 203–209.
[19] Meinkoth, J. and Wahl, G. (1984) Anal. Biochem. 138, 267–284.
[20] Parr, S.R., Barber, D., Greenwood, C., Phillips, B.W. and Melling, J. (1976) Biochem. J. 157, 423–430.
[21] Chater, K.F., Hopwood, D.A., Kieser, T. and Thompson, C.J. (1982) Curr. Top. Microbiol. Immunol. 96, 69–95.
[22] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, Cold Spring Harbour Laboratory, NY.
[23] Birnboim, H.C. and Doly, J. (1979) Nucleic Acids Res. 7, 1513–1523.
[24] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 12, 5463–5467.
[25] Lawton, S.A. and Anthony, C. (1985) J. Gen. Microbiol. 131, 2165–2171.
[26] Tobari, J. (1984) in: Microbial Growth on C_1 Compounds, Proceedings of the 4th International Symposium (Crawford, R.L. and Hanson, R.S. eds) pp.106–112, American Society for Microbiology, Washington.
[27] Ryden, L. and Lundgren, J.O. (1976) Nature 261, 344–346.
[28] Duffaud, G.D., Lehnhardt, S.K., March, P.E. and Inouye, M. (1985) Curr. Top. Membranes Transp. 24, 65–104.
[29] Smeekens, S., De Groot, M., Van Binsbergen, J. and Weisbeek, P. (1985) Nature 317, 456–458.
[30] Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4690.
[31] Watson, M.E.E. (1984) Nucleic Acids Res. 12, 5145–5164.
[32] Von Heijne, G. and Blomberg, C. (1979) Eur. J. Biochem. 97, 175–181.
[33] Husain, M. and Davidson, V.L. (1985) J. Biol. Chem. 260, 14626–14629.
[34] Henry, Y. and Bessières, P. (1984) Biochimie 66, 259–289.