Nitrite reductase from *Pseudomonas aeruginosa*: sequence of the gene and the protein

Maria Chiara Silvestrini*, Cesira L. Galeotti°, Michel Gervais[†], Eugenia Schininà, Donatella Barra, Francesco Bossa and Maurizio Brunori

*Dipartimento di Scienze e Tecnologie Biomediche e Biometria, Università de L'Aquila, località Collemaggio, 67100 L'Aquila, Italy, °Centro Ricerche Sclavo, via Fiorentina 1, 53100 Siena, Italy, [†]Centre de Génétique Moléculaire du CNRS, 91190 Gif sur Yvette, France and Dipartimento di Scienze Biochimiche e Centro di Biologia Molecolare del CNR, Università La Sapienza, P.le A. Moro 5, 00185 Roma, Italy

Received 3 July 1989

The gene coding for nitrite reductase of *Pseudomonas aeruginosa* has been cloned and its sequence determined. The coding region is 1707 bp long and contains information for a polypeptide chain of 568 amino acids. The sequence of the mature protein has been confirmed independently by extensive amino acid sequencing. The amino-terminus of the mature protein is located at Lys-26; the preceding 25 residue long extension shows the features typical of signal peptides. Therefore the enzyme is probably secreted into the periplasmic space. The mature protein is made of 543 amino acid residues and has a molecular mass of 60204 Da. The *c*-heme-binding domain, which contains the only two Cys of the molecule, is located at the amino-terminal region. Analysis of the protein sequence in terms of hydrophobicity profile gives results consistent with the fact that the enzyme is fully water soluble and not membrane bound; the most hydrophilic region appears to correspond to the *c*-heme domain. Secondary structure predictions are in general agreement with previous analysis of circular dichroic data.

Nitrite reductase; Primary structure; Pre-protein; Cytochrome oxidase; (*Pseudomonas aeruginosa*)

1. INTRODUCTION

Pseudomonas nitrite reductase, originally discovered and identified by Horio [1] as an oxidase (ferrocytochrome *c*-551-oxygen oxidoreductase, EC 1.9.3.2), is a soluble redox enzyme synthesized by *Ps. aeruginosa* in the presence of nitrate in the growing medium. The enzyme is a homodimer of 120000 Da, each subunit containing one *c* and one d_1 heme [2–4]. It is able to accomplish the monoelectronic reduction of nitrite to NO as well as the four electronic reduction of oxygen to water, and to interact with two different macromolecular substrates, azurin and cytochrome *c*-551 [5].

A great deal of spectroscopic and functional data have been collected on this protein, providing

Correspondence address: M. Brunori, Dipartimento di Scienze Biochimiche e Centro di Biologia Molecolare del CNR, Università 'La Sapienza', Piazale Aldo Moro 5, 00185 Roma, Italy

information on the interactions of the hemes with the protein moiety and on the mechanism of electron transfer [6,7]. In contrast, the structural information available was relatively poor. In order to partially fill this gap, we have determined the primary structure of the enzyme both by recombinant DNA technology and classical protein sequencing techniques.

Here we report the results of this combined approach, i.e. the complete sequence of nitrite reductase gene confirmed by independent sequence of the protein. Relevant information obtained by protein and DNA sequences analysis is presented and briefly discussed.

2. MATERIALS AND METHODS

Nitrite reductase was purified from *Ps. aeruginosa* (NCTC 6750) according to Parr et al. [8]. The protein was more than

95% pure on the basis of SDS-PAGE [9] and spectroscopic indexes [8].

2.1. Recombinant DNA techniques

Chromosomal DNA from *Ps. aeruginosa* (NCTC 6750) was purified from 10 g of cell paste according to Chater [10] as modified by Canters [11]. A DNA library was constructed in λgt11 from partially *Eco*RI-digested DNA according to Huyhn et al. [12]. Amplification of the library in *E. coli* Y1090 and subsequent screening with labeled oligonucleotides were performed as described by Huyhn et al. for λgt10 libraries [12]. Filter hybridization conditions were 2 × SSC, 55°C. The positive insert was subcloned in the plasmid pEMBL 18 [13]. Purification of λgt11 DNA, pEMBL 18 DNA, restriction enzyme digestions, Southern blotting and labeling of oligonucleotides were performed following standard protocols [14].

Oligonucleotides for screening the library and for DNA sequencing were kindly synthesized by S. Ricci at Centro Ricerche Sclavo (Siena, Italy).

DNA sequencing was carried out according to Sanger et al. [15] using synthetic 20-mer oligonucleotides as primers. In order to resolve compressions due to the very high GC content of *Ps. aeruginosa* DNA, sequence reactions containing dIPT instead of dGTP were run in parallel to standard reactions [16].

2.2. Protein sequencing techniques

9.5 mg of nitrite reductase were carboxymethylated according to Barra et al. [17] and cleaved with CNBr [18]. The digest was fractionated by gel filtration on a Sephadex G50 (superfine) column (2 cm × 130 cm) equilibrated and eluted with 10% acetic acid. Where necessary, aliquots of peptides were subjected to secondary digestion with trypsin. Isolation of tryptic peptides was carried out by high-performance liquid chromatography on a macroporous reverse-phase column (Aquapore RP-300, 4.6 mm × 250 mm, 10 μm, Brownlee Labs) with gradients of 0–70% acetonitrile in 0.2% trifluoroacetic acid. Elution of peptides was monitored on a Beckman 165 spectrophotometer at 220 and 280 nm. Sequence analyses of the intact protein and of the purified peptides were performed on an Applied Biosystems model 470A gas-phase protein sequencer equipped with an Applied Biosystems model 120A PTH analyzer for the on-line detection of phenylthiohydantoin amino acids. Samples were loaded onto trifluoroacetic-treated glass-fiber filters, coated with polybrene and prewashed according to the manufacturer's instructions.

3. RESULTS

3.1. Isolation of the nitrite reductase gene

Five oligonucleotides (table 1) were synthesized on the basis of the sequence of a suitable portion of nitrite reductase *c*-heme-binding peptide [19], i.e. the stretch Phe-Asn-Glu-Ala-Lys-Gln-Ile-Tyr-Phe-Gln. These were used as a mixture to screen the *Ps. aeruginosa* DNA library. The oligonucleotide sequences were chosen on the basis of the codon usage frequencies of 8 *Ps. aeruginosa* genes (see the

Table 1

Synthetic oligonucleotides used for screening of the *Ps. aeruginosa* DNA library

NR1	5' GGAAGTAGATCTGCTTGGCCTCGTTGAA 3'
NR2	5' GGAAGTAGATCTGCTTCGCCTCGTTGAA 3'
NR3	5' GGAAGTAGATCTGCTTGGCTTCGTTGAA 3'
NR4	5' GGAAGTAGATCTGCTTCGCTTCGTTGAA 3'
NR5	5' GGAAATAGATCTGCTTGGCTTCGTTGAA 3'

The table shows the sequence of the five oligonucleotides used for screening the *Ps. aeruginosa* DNA library. The sequences are complementary to five of the possible sequences coding for the peptide Phe-Asn-Glu-Ala-Lys-Gln-Ile-Tyr-Phe-Gln in the *c*-heme-binding region ([19] and positions 61–70 of the complete sequence in fig.2). They represent the five most probable ones on the basis of *Ps. aeruginosa* codon usage frequencies. These frequencies, calculated from the sequences of azurin [11], esotoxin A [20], mercuric reductase [21], metapyrocatechase [22], carboxypeptidase G2 [23], phosphomannosoisomerase [24], GDP mannose dehydrogenase [25], and acylase [26] are as follows: Glu, GAG 65%; Phe, TTC 96%; Asn, AAC 91%; Ala, GCC 55% and GCG 36%; Lys, AAG 89%; Gln, CAG 87%; Ile, ATC 91%, Tyr, TAC 75%. They do not differ significantly from those calculated from the nitrite reductase gene. The probability of success for each of the selected nucleotides, calculated as composed probability from the above frequencies, is: NR1, 15.8%; NR2, 10.4%; NR3, 8.5%; NR4, 5.6%; NR5, 5.3% giving a total probability of success of 46.6%

legend of table 1). The screening gave one positive clone out of approx. 5000 plaques.

*Eco*RI digestion of the phage purified from the positive clone produced the arms of the vector and two fragments of approx. 3.5 and 2 kb. In Southern blot analysis of *Eco*RI-digested DNA, using the mixture of five oligonucleotides as a probe, only the 3.5 kb fragment gave a positive signal. A stronger signal was observed using a mixture of probes 1 and 2 with respect to a mixture of probes 3,4 and 5. Sequencing of the gene showed that the sequence of probe 1 was the correct one. The 3.5 kb fragment was subcloned into the *Eco*RI site of the plasmid pEMBL 18.

A restriction map of the subcloned fragment was determined for several enzymes and the hybridization site for the oligonucleotide probes was shown by Southern blotting to be located nearby (approx. 300 bp) the universal primer site of the plasmid (fig.1).

3.2. Nucleotide sequence

Most of the sequencing was carried out on the *Eco*RI pEMBL 18 subclone. A *Sal*I pEMBL 18
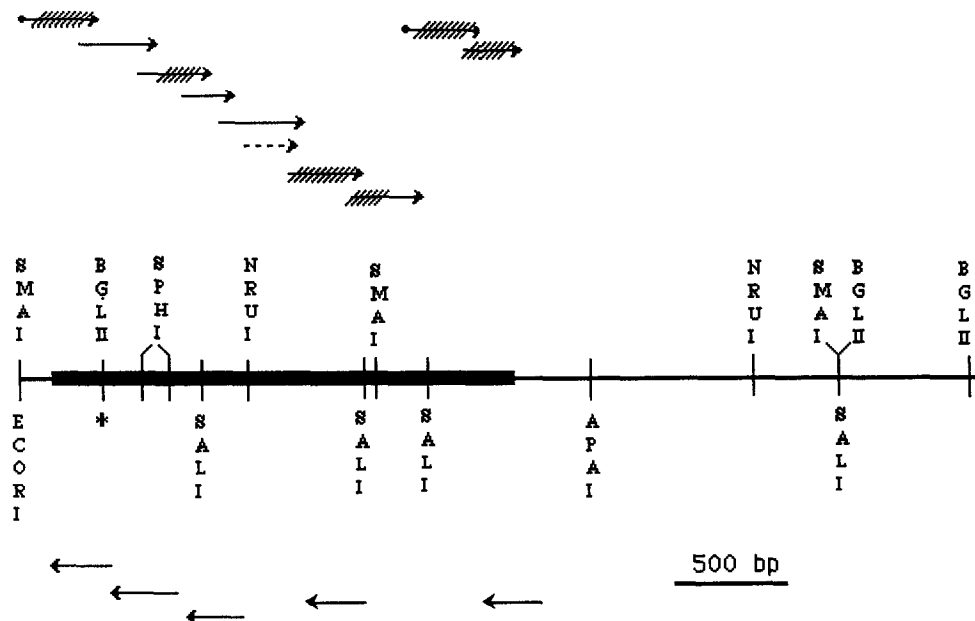
Fig.1. Restriction map and sequencing strategy for the *Ps. aeruginosa* nitrite reductase gene. The 3.5 kb *Eco*RI fragment subcloned in pEMBL 18 is shown with major restriction sites. The wide bar indicates the ORF. (∗) The position of probe 1 (table 1). Sequenced fragments are shown by rightward arrows (sense strand) and leftward arrows (complementary strand). The M13mp18 universal primer (arrows starting with a dot) and several synthetic oligonucleotides were used as primers. Primers sequences were all based on DNA sequencing data except for one case (dashed arrow), in which the primer was synthesized on the basis of the amino-terminal sequence of a peptide obtained from CNBr protein digestion, containing a suitable sequence, i.e. Met-Ile-Asp-Leu-Trp-Ala-Lys (positions 229–235 in fig.2). The amino acid sequence was translated as 5′-ATGATCGACCTGTGGGCCAA-3′ following the criteria described in table 1 and contains one mismatch with the actual DNA sequence.

subclone containing the *Sal*I fragment of 1.5 kb (fig.1) in the same orientation as the *Eco*RI subclone was also obtained and used for sequencing.

The sequencing strategy is outlined in fig.1. The nucleotide sequence, shown in fig.2, contains one open reading frame (ORF) which starts with an ATG at position 82 and ends with a TGA at position 1788. Translation of this ORF gives a protein of 568 amino acids (fig.2). Analysis of the DNA composition shows that the nitrite reductase gene is very rich in GC bases (64%). A codon usage analysis (not shown) confirms the very high preference for codons ending with G or C except in the case of Glu and His.

3.3. Protein sequence

Automated Edman degradation of the native protein was performed for 45 cycles. The resulting sequence corresponds exactly to that deduced from the DNA sequence at positions 26 to 70. Thus

nitrite reductase is synthesized in vivo as a preprotein, with a 25 residue long signal peptide.

A number of CNBr and tryptic fragments of the protein were also sequenced; their structure is indicated in fig.2. Information from the protein sequence accounts for 454 residues, that is 84% of the structure of the mature protein.

The amino acid sequence has been analyzed with available predictive algorithms [27–29]; the results of hydrophobicity analysis and of secondary structure predictions are shown in fig.3.

4. DISCUSSION

We have determined the complete primary structure of *Ps. aeruginosa* nitrite reductase with a combined approach of recombinant DNA technology and classical biochemical techniques. Knowledge on the structure of the nitrite reductase gene is the first step for further investigations on structure-

```
GAATTCCCGGGAGTTCCCGACGCAGCCACCCCCAAAACACTGCTAAGGGAGCGCCTCGCAGGGCTCCTGAGGAGATAGACC ATG CCA TTT GGC AAG CCA CTG GTG GGC ACC TTG CTC GCC TCG CTG   126
                                                                                 Met Pro Phe Gly Lys Pro Leu Val Gly Thr Leu Leu Ala Ser Leu    15

ACG CTG CTG GGC CTG GCC ACC GCT CAC GCC AAG GAC GAC ATG AAA GCC GCC GAG CAA TAC CAG GGT GCC GCT TCC GCC GTC GAT CCC GCT CAC GTG GTG CGC ACC AAC   234
Thr Leu Leu Gly Leu Ala Thr Ala His Ala Lys Asp Asp Met Lys Ala Ala Glu Gln Tyr Gln Gly Ala Ala Ser Ala Val Asp Pro Ala His Val Val Arg Thr Asn    51

GGC GCT CCC GAC ATG AGT GAA AGC GAG TTC AAC GAG GCC AAG CAG ATC TAC TTC CAA CGC TGC GCC GGT TGC CAC GGC GTC CTG CGC AAG GGC GCC ACC GGC AAG CCG   342
Gly Ala Pro Asp Met Ser Glu Ser Glu Phe Asn Glu Ala Lys Gln Ile Tyr Phe Gln Arg Cys Ala Gly Cys His Gly Val Leu Arg Lys Gly Ala Thr Gly Lys Pro    87

CTG ACC CCG GAC ATC ACC CAG CAA CGC GGC CAG CAA TAC CTG GAA GCG CTG ATC ACC TAC GGC ACC CCG CTG GGC ATG CCG AAC TGG GGC AGC TCC GGC GAG CTG AGC   450
Leu Thr Pro Asp Ile Thr Gln Gln Arg Gly Gln Gln Tyr Leu Glu Ala Leu Ile Thr Tyr Gly Thr Pro Leu Gly Met Pro Asn Trp Gly Ser Ser Gly Glu Leu Ser   123

AAG GAA CAG ATC ACC CTG ATG GCC AAG TAC ATC CAG CAC ACC CCG CCG CAA CCG CCG GAG TGG GGC ATG CCG GAG ATG CGC GAA TCG TGG AAG GTG CTG GTG AAG CCG   558
Lys Glu Gln Ile Thr Leu Met Ala Lys Tyr Ile Gln His Thr Pro Pro Gln Pro Pro Glu Trp Gly Met Pro Glu Met Arg Glu Ser Trp Lys Val Leu Val Lys Pro   159

GAG GAC CGG CCG AAG AAA CAG CTC AAC GAC CTC GAC CTG CCC AAC CTG TTC TCG GTG ACC CTG CGC GAC GCC GGG CAG ATC GCC CTG GTC GAC GGC GAC AGC AAA AAG   666
Glu Asp Arg Pro Lys Lys Gln Leu Asn Asp Leu Asp Leu Pro Asn Leu Phe Ser Val Thr Leu Arg Asp Ala Gly Gln Ile Ala Leu Val Asp Gly Asp Ser Lys Lys   195

ATC GTC AAG GTC ATC GAT ACC GGC TAT GCC GTG CAT ATC TCG CGG ATG TCC GCT TCC GGC CGC TAC CTG CTG GTG ATC GGC CGC GAC GCG CGG ATC GAC ATG ATC GAC   774
Ile Val Lys Val Ile Asp Thr Gly Tyr Ala Val His Ile Ser Arg Met Ser Ala Ser Gly Arg Tyr Leu Leu Val Ile Gly Arg Asp Ala Arg Ile Asp Met Ile Asp   231

CTG TGG GCC AAG GAG CCG ACC AAG GTC GCC GAG ATC AAG ATC GGC ATC GAG GCG CGC TCG GTG GAA AGC TCC AAG TTC AAG GGC TAC GAG GAC CGC TAC ACC ATC GCC   882
Leu Trp Ala Lys Glu Pro Thr Lys Val Ala Glu Ile Lys Ile Gly Ile Glu Ala Arg Ser Val Glu Ser Ser Lys Phe Lys Gly Tyr Glu Asp Arg Tyr Thr Ile Ala   267

GGC GCC TAC TGG CCG CCG CAG TTC GCG ATC ATG GAC GGC GAG ACC CTG GAA CCG AAG CAG ATC GTC TCC ACC CGC GGC ATG ACC GTA GAC ACC CAG ACC TAC CAC CCG   990
Gly Ala Tyr Trp Pro Pro Gln Phe Ala Ile Met Asp Gly Glu Thr Leu Glu Pro Lys Gln Ile Val Ser Thr Arg Gly Met Thr Val Asp Thr Gln Thr Tyr His Pro   303

GAA CCG CGC GTG GCG GCG ATC ATC GCC TCC CAC GAG CAC CCC GAG TTC ATC GTC AAC GTG AAG GAG ACC GGC AAG GTC CTG CTG GTC AAC TAC AAG GAT ATC GAC AAC   1098
Glu Pro Arg Val Ala Ala Ile Ile Ala Ser His Glu His Pro Glu Phe Ile Val Asn Val Lys Glu Thr Gly Lys Val Leu Leu Val Asn Tyr Lys Asp Ile Asp Asn   339

CTC ACC GTC ACC AGC ATC GGT GCG GCG CCG TTC CTC CAC GAC GGC GGC TGG GAC AGC AGC CAC CGC TAC TTC ATG ACC GCC GCC AAC AAC TCC AAC AAG GTT GCC GTG   1206
Leu Thr Val Thr Ser Ile Gly Ala Ala Pro Phe Leu His Asp Gly Gly Trp Asp Ser Ser His Arg Tyr Phe Met Thr Ala Ala Asn Asn Ser Asn Lys Val Ala Val   375

ATC GAC TCC AAG GAC CGT CGC CTG TCG GCC CTG GTC GAC GTC GGC AAG ACC CCG CAC CCG GGG CGT GGC GCC AAC TTC GTG CAT CCC AAG TAC GGC CCG GTG TGG AGC   1314
Ile Asp Ser Lys Asp Arg Arg Leu Ser Ala Leu Val Asp Val Gly Lys Thr Pro His Pro Gly Arg Gly Ala Asn Phe Val His Pro Lys Tyr Gly Pro Val Trp Ser   411

ACC AGC CAC CTG GGC GAC GGC AGC ATC TCG CTG ATC GGC ACC GAT CCG AAG AAC CAT CCG CAG TAC GCC TGG AAG AAA GTC GCC GAA CTA CAG GGC CAG GGC GGC GGC   1422
Thr Ser His Leu Gly Asp Gly Ser Ile Ser Leu Ile Gly Thr Asp Pro Lys Asn His Pro Gln Tyr Ala Trp Lys Lys Val Ala Glu Leu Gln Gly Gln Gly Gly Gly   447

TCG CTG TTC ATC AAG ACC CAT CCG AAG TCC TCG CAC CTC TAC GTC GAC ACC ACC TTC AAC CCC GAC GCC AGG ATC AGC CAG AGC GTC GCG GTG TTC GAC CTG AAG AAC   1530
Ser Leu Phe Ile Lys Thr His Pro Lys Ser Ser His Leu Tyr Val Asp Thr Thr Phe Asn Pro Asp Ala Arg Ile Ser Gln Ser Val Ala Val Phe Asp Leu Lys Asn   483

CTC GAC GCC AAG TAC CAG GTG CTG CCG ATC GCC GAA TGG GCC GAT CTC GGC GAA GGC GCC AAG CGG GTG GTG CAG CCC GAG TAC AAC AAG CGC GGC GAT GAA GTC TGG   1638
Leu Asp Ala Lys Tyr Gln Val Leu Pro Ile Ala Glu Trp Ala Asp Leu Gly Glu Gly Ala Lys Arg Val Val Gln Pro Glu Tyr Asn Lys Arg Gly Asp Glu Val Trp   519

TTC TCG GTG TGG AAC GGC AAG AAC GAC AGC TCC GCG CTG GTG GTG GTG GAC GAC AAG ACC CTG AAG CTC AAG GCC GTG GTC AAG GAC CCG CGG CTG ATC ACC CCG ACC   1746
Phe Ser Val Trp Asn Gly Lys Asn Asp Ser Ser Ala Leu Val Val Val Asp Asp Lys Thr Leu Lys Leu Lys Ala Val Val Lys Asp Pro Arg Leu Ile Thr Pro Thr   555

GGT AAG TTC AAC GTC TAC AAC ACC CAG CAC GAC GTG TAC TGA GACCCGCGTGCGGGGCACGCCCCGCACGCTCCCCCCTACGAGGAACCGTGATGAAACCGTACGCACTGCTTTCGCTGCTCGCCA   1872
Gly Lys Phe Asn Val Tyr Asn Thr Gln His Asp Val Tyr END
```

Fig.2. The complete nucleotide sequence of the *Ps. aeruginosa* nitrite reductase gene and the amino acid sequence of the protein. The nucleotide sequence is written in the 5′–3′ direction of the coding strand. The deduced amino acid sequence is shown below. The putative ribosome-binding site is indicated in a box. The arrows underline a palindromic region probably important for transcription termination. Peptides for which the sequence has already been determined by amino acid sequencing are underlined by a continuous line. The heme-binding peptide previously sequenced [19] is underlined with a dashed line.
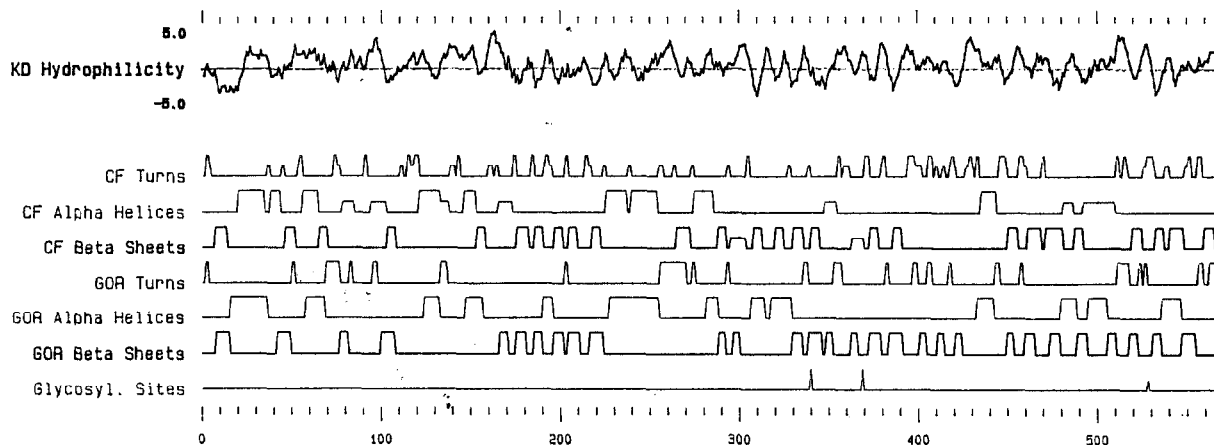
Fig.3. Hydrophilicity profile and secondary structure predictions for *Ps. aeruginosa* nitrite reductase. The hydrophilicity profile of the nitrite reductase pre-protein was calculated by the method of Kyte and Doolittle ($n = 9$) [27]. Secondary structure predictions according to Chou and Fasman [28] (CF) and Garnier et al. [29] (GOR) are reported.

function relationships by site-directed mutagenesis. Moreover, studies on regulation of gene expression, which are particularly attractive for this inducible system, become feasible.

The DNA sequence of the nitrite reductase gene is particularly rich in GC (64%), a prominent feature of *Ps. aeruginosa* genes [20–26]. This caused localized artifacts in sequencing gels. Extensive protein sequencing was useful in identifying these artifacts and, combined with the DNA sequencing strategy outlined in fig.1, in solving these ambiguities.

No differences in base composition are observable between coding and noncoding regions, but this may be due to the small fraction of noncoding region sequenced. The *Eco*RI site is rather near to the start codon (81 bp); thus the promoter region with high AT content is probably missing in this clone. A putative ribosome-binding region (box in fig.2) is present just upstream of the ATG codon, as expected [30]. Downstream of the termination codon, a potential transcription terminator sequence with dyad symmetry is present (fig.2) [31].

Direct sequencing of the protein not only proves the identity of the cloned gene, but also provides information on the structure of the mature protein. Comparison of the protein sequence deduced by the DNA sequence with the amino-terminus of the native protein shows that nitrite reductase is synthesized as a pre-protein, with a 25 residue long leader peptide. Analysis of this amino-terminal ex-

tension (pos. 1–25 in fig.2) shows that it has all the characteristic features of the known presecretory signal peptides [32,33], i.e.: (i) a hydrophobic core of correct length (pos. 7–12), very rich in Leu, starting with a secondary structure disrupting residue (Pro-6) just after a positively charged residue (Lys-5); (ii) an Ala-X-Ala sequence preceding the peptidase cleavage site. From the physiological view point, the presence of a signal peptide suggests that the protein is secreted in the periplasmic space. Previous results [34] indicated that not only nitrite reductase, but also azurin and cytochrome *c*-551 are located in the periplasmic space. However, the enzyme was shown to be bound at the inner surface of the cytoplasmic membrane by labeling with ferritin-conjugated antibodies [35]. The existence of a leader peptide supports the periplasmic location of nitrite reductase and is perfectly in line with the finding that also the primary product of azurin gene has a signal peptide [11].

The mature protein is 543 amino acids long; the molecular mass of 60 204 Da supports previous estimates (58–65 000 Da for the subunit and 119–130 000 for the native dimer) based on iron content, SDS-PAGE, sedimentation equilibrium and diffusion measurements [2–4]. The amino acid composition is in good agreement with previous measurements [36,37]. The *c*-heme-binding peptide, with a sequence identical to that already published [19], is located nearby the amino-terminus of the protein: the two Cys covalently

bound to the protoporphyrin vinyl groups are at positions 75 and 78, and the Met proposed to be the sixth heme ligand, on the basis of amino acid alignment with other bacterial cytochromes c [6], is at position 113.

Analysis of the protein sequence in terms of hydrophobicity and secondary structure provides preliminary information of some interest. The c-heme-binding domain is clearly identified, and appears to contain several α-helix stretches; on the other hand most of the β-sheets are present in the last 3/4 of the sequence. These predictions are in line with the analysis of the far ultraviolet circular dichroism of the protein, previously reported [38]. Given that the dimeric structure is mainly stabilized by hydrophobic interactions [39], the less hydrophilic regions (which are also more prominent after residue 150) may be involved in the monomer-monomer contact.

More extensive analysis of the structure, by a combined use of proteolytic fragmentation and predictive methods, is presently underway.

REFERENCES

[1] Horio, T., Higashi, T., Matsubara, H., Kusai, K., Nakai, M. and Okunuki, K. (1958) Biochim. Biophys. Acta 29, 297–302.

[2] Kuronen, T. and Ellfolk, N. (1972) Biochim. Biophys. Acta 275, 308–318.

[3] Gudat, J.C., Singh, J. and Wharton, D.C. (1973) Biochem. Biophys. Res. Commun. 292, 376–390.

[4] Silvestrini, M.C., Colosimo, A., Brunori, M., Walsh, T.A., Barber, D. and Greenwood, C. (1979) Biochem. J. 183, 701–709.

[5] Yamanaka, T., Ota, A. and Okunuki, K. (1961) Biochim. Biophys. Acta 53, 294–308.

[6] Meyer, T. and Kamen, M.D. (1982) Adv. Prot. Chem. 35, 162–170.

[7] Pettigrew, W.P. and Moore, G.R. (1987) Cytochromes c: Biological Aspects, pp. 161–168, Springer-Verlag.

[8] Parr, S.R., Barber, D., Greenwood, C., Phillips, B.W. and Melling, J. (1976) Biochem. J. 157, 423–430.

[9] Laemmli, U.K. (1976) Nature 227, 680–682.

[10] Chater, K.F., Hopwood, D.A., Kieser, T. and Thompson, C.J. (1982) Curr. Top. Microbiol. Immunol. 96, 69–95.

[11] Canters, G.W. (1987) FEBS Lett. 212, 168–172.

[12] Huynh, T.V., Young, R.A. and Davis, R.W. (1985) in: DNA Cloning (vol I): A Practical Approach (Glover, D.M. ed.) pp. 49–78, IRL Press, Oxford.

[13] Dente, L., Sollazzo, M., Baldari, C., Cesareni, G. and Cortese, R. (1985) in: DNA Cloning (vol. I): A Practical Approach (Glover, D.M. ed.) pp. 101–135, IRL Press, Oxford.

[14] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

[15] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.

[16] Mills, D.R. and Kramer, F.R. (1979) Proc. Natl. Acad. Sci. USA 76, 2232–2235.

[17] Barra, D., Schininà, E., Simmaco, M., Bannister, J.V., Bannister, W.H., Rotilio, G. and Bossa, F. (1984) J. Biol. Chem. 259, 12595–12601.

[18] Gross, E. and Witkop, B. (1961) J. Am. Chem. Soc. 83, 1510–1511.

[19] Kalkkinen, N. and Ellfolk, N. (1978) IUPAC 11th Int. Symp. Chem. Nat. Prod., pp. 79–82.

[20] Gray, G.L., Smith, D.H., Balbridge, J.S., Harkins, R.N., Vasi, M.L., Chen, E.Y. and Heyneker, H.L. (1984) Proc. Natl. Acad. Sci. USA 81, 2645–2649.

[21] Brown, N.L., Ford, S.J., Pridmore, R.D. and Fritzinger, D.C. (1983) Biochemistry 22, 4089–4095.

[22] Nakai, C., Kagamiyama, H., Nozaki, M., Nakazawa, T., Inouye, S., Ebina, Y. and Nakazawa, A. (1983) J. Biol. Chem. 258, 2923–2928.

[23] Minton, N.P., Atkinson, T. and Sherwood, R.F. (1983) J. Bacteriol. 156, 1222–1227.

[24] Darzins, A., Frantz, B., Vanags, R.I. and Chakrabarty, A.M. (1986) Gene 42, 293–302.

[25] Deretic, V., Gill, J.F. and Chakrabarty, A.M. (1987) Nucleic Acids Res. 15, 4567–4581.

[26] Matsuda, A. and Komatsu, K.I. (1985) J. Bacteriol. 163, 1222–1228.

[27] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.

[28] Chou, P.Y. and Fasman, G.D. (1978) Adv. Enzymol. 47, 45–148.

[29] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) J. Mol. Biol. 120, 97–120.

[30] Shine, J. and Dalgarno, L. (1975) Nature 254, 34–38.

[31] Adhya, S. and Gottesman, M. (1978) Ann. Rev. Biochem. 47, 967–996.

[32] Von Heijne, G. (1983) Eur. J. Biochem. 133, 17–21.

[33] Perlman, D. and Halvorson, H.O. (1983) J. Mol. Biol. 167, 391–409.

[34] Wood, P.M. (1978) FEBS Lett. 92, 214–218.

[35] Saraste, M. and Kuronen, T. (1978) Biochim. Biophys. Acta 513, 117–131.

[36] Nagata, Y., Yamanaka, T. and Okunuki, K. (1970) Biochim. Biophys. Acta 221, 668–671.

[37] Silvestrini, M.C., Citro, G., Colosimo, A., Chersi, A., Zito, R. and Brunori, M. (1983) Anal. Biochem. 129, 318–325.

[38] Tordi, M.G., Silvestrini, M.C., Colosimo, A., Provencher, S. and Brunori, M. (1984) Biochem. J. 218, 907–912.

[39] Kuronen, T., Saraste, M. and Ellfolk, N. (1975) Biochim. Biophys. Acta 393, 48–54.