Immuno-purification of a dimeric subcomplex of the respiratory NAHD-CoQ reductase of *Rhodobacter capsulatus* equivalent to the FP fraction of the mitochondrial complex I

Hervé Duborjal[a], Alain Dupuis[a,*], Agnès Chapel[b], Silvie Kieffer[b], Joël Lunardi[a], Jean-Paul Issartel[a]

[a]*Laboratoire de Bioénergétique Cellulaire et Pathologique, EA 2019 UJF, DBMS, CEA Grenoble, 38054 Grenoble Cedex 9, France*
[b]*Laboratoire de Chimie des Protéines, DBMS, CEA Grenoble, 38054 Grenoble Cedex 9, France*

Abstract The *Rhodobacter capsulatus* genes encoding the NUOE and NUOF subunits, equivalent to the 24 kDa and 51 kDa subunits of the mammalian mitochondrial complex I, have been sequenced. According to the nucleotide sequence, the NUOE subunit is 389 amino acids long and has a molecular mass of 41.3 kDa. In comparison to the mitochondrial equivalent subunit, NUOE is extended at the C terminus by more than 150 amino acids. The NUOF subunit is 431 amino acids long and has a molecular mass of 47.1 kDa. A subcomplex containing both the NUOE and NUOF subunits was extracted by detergent treatment of *R. capsulatus* membranes and immuno-purified. This subcomplex is homologous to the mitochondrial FP fragment. Mass spectrometry after trypsin treatment of the NUOE subunit validates the atypical primary structure deduced from the sequence of the gene.

© 1997 Federation of European Biochemical Societies.

Key words: Rhodobacter capsulatus; Complex I; Mitochondrion; Gene sequence; Immunopurification; MALDI mass spectrometry

1. Introduction

NADH-CoQ reductase (or complex I) is the first component of the bacterial or mitochondrial respiratory chains. This enzyme catalyzes the oxidation of NADH and the reduction of ubiquinone with the concomitant ejection of 4 protons outside the bacterial cell or mitochondria [1,2]. In bovine mitochondria, structural analysis has revealed this complex to be composed of 43 different subunits. A number of cofactors (one FMN and 4–6 iron-sulfur clusters) are key components in the redox process. The precise assignment of these bound cofactors to the different subunits is still a matter of debate. Years ago, the subfractionation of the mitochondrial complex I led to the identification of three different subcomplexes called HP, IP and FP with the following features [3,4]. The HP subcomplex (hydrophobic proteins) is a non-soluble multimeric aggregate of hydrophobic subunits. It is partly built up of the 7 subunits encoded by the mitochondrial episome (also called ND subunits: ND1 through ND6 and ND4L). This hydrophobic fraction plays an important role in the anchoring

of complex I to the internal mitochondrial membrane and also in the proton channeling through the membrane. The second subcomplex, called IP fragment (iron-sulfur proteins), is a soluble domain made up of numerous subunits and contains some of the iron-sulfur clusters of complex I. The third subcomplex has been called FP fragment (flavoprotein) as it was proved to contain the FMN molecule and iron-sulfur clusters [5,6]. This soluble domain, when purified from bovine complex I, is composed of three subunits with molecular masses of 51, 24 and 9 kDa and displays NADH-ferricyanate oxidoreductase activity. The 51 kDa subunit is believed to be the catalytic subunit since it binds NADH [7,8]. One 4Fe-4S cluster and one FMN are probably liganded to this subunit. The second iron-sulfur cluster borne by the FP fragment is a 2Fe-2S cluster and is localized to the 24 kDa subunit. Complex I from bacteria, a more simple model of complex I, is being studied in different laboratories. In *Escherichia coli* and *Paracoccus denitrificans,* complex I is believed to be an assembly of at least 14 different subunits [9,10]. The genes encoding these subunits are clustered in one operon (so-called *nuo* for NAHD-ubiquinone oxidoreductase in *E. coli* and *nqo* for NADH-quinone oxidoreductase in *P. denitrificans*). Primary structure analysis has indicated that the 14 bacterial subunits are homologous to 14 subunits which are part of the bovine mitochondrial complex I. Seven bacterial subunits are equivalent to the mitochondrial ND subunits and seven others are the counterparts of the 75 kDa, 51 kDa, 49 kDa, 30 kDa, 24 kDa, 23 kDa and 20 kDa subunits. Hence, two of the bacterial subunits are highly similar to the two main subunits which compose the mitochondrial FP fragment (51 kDa and 24 kDa subunits). The third subunit of the mitochondrial FP fragment called 9 kDa has no counterpart amongst the subunits encoded by the bacterial operons. The present report shows that an FP-like subcomplex made up of two subunits homologous to the mitochondrial 51 kDa and 24 kDa subunits can be immuno-purified from the *R. capsulatus* membranes. Strikingly enough, the *R. capsulatus* NUOE subunit has a very long C terminus extension which is not found in the equivalent subunits of the different bacterial and mitochondrial complexes I.

2. Materials and methods

An *R. capsulatus* B10 genomic DNA library constructed in lambda was screened with a bovine cDNA fragment coding for the 51 kDa subunit. Inserts from the positive clones were subcloned in plasmids and sequenced as described previously [11]. A 12 amino acid peptide called peptide PNE corresponding to the N terminus of the NUOE subunit was synthesized on an Applied Biosystems 432A peptide syn-

*Corresponding author. Fax: (33) 4 76 88 51 87.
E-mail: Dupuis@aup.ceng.cea.fr

Abbreviations: NUO, NADH-ubiquinone oxidoreductase; FP, flavoprotein fraction; MALDI, matrix-assisted laser desorption/ionization; PNE, synthetic peptide corresponding to the NUOE subunit N-terminus

thesizer (the PNE sequence was MLRRLHATQPDSC, in which the C terminus cysteine was an additional residue added for cross-linking to BSA). The HPLC purified peptide was cross-linked to maleimide-activated BSA and the final conjugate was used to raise antibodies in rabbits. Cytoplasmic membranes from aerobically grown wild type *R. capsulatus* B10 strain were prepared as follows. Cells were grown in the dark in RCV medium [12] containing 30 mM lactate as a carbon source and harvested by centrifugation. They were resuspended in one tenth the initial volume in buffer A (50 mM HEPES, pH 7.5, 1 mM EDTA, 1 mM β-mercaptoethanol, 2 μg/ml aprotinin, 1 μg/ml antipain, 1 μg/ml leupeptin, 1 μg/ml chymostatin, 0.5 μg/ml pepstatin) and disrupted by passing through a cell disrupter (Constant Cell, War-

wick, UK) at a pressure of 1 kbar. Subcomplex extraction was performed by treatment of the membranes with a stripping solution composed of 50 mM HEPES pH 7.5, 20 mM NaCl, 1 mM EDTA and 2% Glucopon (Glucopon 215 CS UP is a proprietary detergent from Henkel, France). After centrifugation (15 min, 90 000 rpm, 4°C) the supernatant was diluted five-fold in phosphate saline buffer and loaded onto a 1 ml immunoaffinity column of immobilized anti-PNE antibodies. The immunoaffinity column was generated essentially as described by Haines et al. [13]. Very briefly, the anti-PNE antibodies were immobilized on a protein A crosslinked Sepharose CL6B (Pharmacia) then chemically crosslinked with dimethylpimelimidate. The Glucopon solubilized *R. capsulatus* membrane extract was loaded

```
1      atcgatgccgtcgccaagggacaccaactggcggatgtgtccgccatcatcggcacgatggatgtcgtgttcggggagattgaccgctg ATG CTG  95
1                                                                                          NUOE  M   L    2
96     CGC CGC TTG CAC GCC ACC CAA CCC GAC AGC TTC GCC TTT ACC CCC GCC AAC CGC GCC TGG GCG GAA GCG CAG ATC  170
3       R   R   L   H   A   T   Q   P   D   S   F   A   F   T   P   A   N   R   A   W   A   E   A   Q   I    27
171    ACC AAA TAC CCC GAG GGC CGT CAG GCC TCG GCG GTG ATC CCG CTC TTG TTC CGC GCG CAG GAA CAG GAA GGC TGG  245
28      T   K   Y   P   E   G   R   Q   A   S   A   V   I   P   L   L   F   R   A   Q   E   Q   E   G   W    52
246    CTG ACG AAG CCC GCG ATC GAA TAC GTG GCC GAC CTG CTC GAC ATG CCC TAT ATC CGC GTG CTC GAA GTG GCG TCC  320
53      L   T   K   P   A   I   E   Y   V   A   D   L   L   D   M   P   Y   I   R   V   L   E   V   A   S    77
321    TTC TAC TTC ATG TTC CAG CTG CAG CCC ACG GGC TCG GTC GCG CAT ATT CAG GTC TGC GGC ACC ACC AGC TGC ATG  395
78      F   Y   F   M   F   Q   L   Q   P   T   G   S   V   A   H   I   Q   V   C   G   T   T   S   C   M    102
396    ATC ATG GGC TCG GAA AAC CTG ATC GAG GTC TGC AAG CGC AAG ATC TCG CAT CAT CCG CAT GCG CTT TCC GCC GAT  470
103     I   M   G   S   E   N   L   I   E   V   C   K   R   K   I   S   H   H   P   H   A   L   S   A   D    127
471    GGC AAG TTC AGC TGG GAA GAG GTC GAA TGT CTG GGC GCC TGT GCC AAT GCG CCG ATG GCC CAG ATC GGC AAG GAT  545
128     G   K   F   S   W   E   E   V   E   C   L   G   A   C   A   N   A   P   M   A   Q   I   G   K   D    152
546    TAC TTC GAA GAC CTG ACC GAG GCG GGG CTG GAG CGG CTG ATC GAC GAT CTG GCG GCG GGC AAG GCG CCG GTG CCG  620
153     Y   F   E   D   L   T   E   A   G   L   E   R   L   I   D   D   L   A   A   G   K   A   P   V   P    177
621    GGG TCG GAA ATC GGG CGC TTC GGG GCG GAA CCG GCC ACC GGC CTG ACC AGC TGC ACG ACG ACC TCC GGC GCG CGC  695
178     G   S   E   I   G   R   F   G   A   E   P   A   T   G   L   T   S   C   T   T   T   S   G   A   R    202
696    GAG GCG CAG AAC GGC TCG GTC GCG CGG GCG GTT GCG GTG GGC GAT GGC ACG CGC CGG ATC GAC GGG ACC GAG GTG  770
203     E   A   Q   N   G   S   V   A   R   A   V   A   V   G   D   G   T   R   R   I   D   G   T   E   V    227
771    CGG CTT TCC GCC CCA TGG CAG GTC TCG GGG CAG ATG GGC GCG ATC GGT CTG ATG ACG GCG CGG CCG CGC TCG GCG  845
228     R   L   S   A   P   W   Q   V   S   G   Q   M   G   A   I   G   L   M   T   A   R   P   R   S   A    252
846    GTG CAG CAG CAG GCG GCG GCC GAG GCG GCG AAA GCG GCG GCG GCA GAG TCG GCC GGG GGG ACG TCA ACC CCG  920
253     V   Q   Q   Q   A   A   A   E   A   A   K   A   A   A   A   A   E   S   A   G   G   T   S   T   P    277
921    GCT GCA CCG GTG GCC GGG GCG GCG GGG AAA GCC GCA CAA ACG GCT GAA ACC GCT GCG CAG AAA CCG CTT GCG CTG  995
278     A   A   P   V   A   G   A   A   G   K   A   A   Q   T   A   E   T   A   A   Q   K   P   L   A   L    302
996    GTT GCA CCC CGC GAC GGC AAG GCC GAT GAG CTG CAG AAG ATC GAG AGC ATC GGA CCG GGC CTC GAG GCG CTT CTG  1070
303     V   A   P   R   D   G   K   A   D   E   L   Q   K   I   E   S   I   G   P   A   L   E   A   L   L    327
1071   CAC GAT CTG GGC ATC TTC CAT TTC GAC CAG ATC GCC GCC TGG GGC GCG GGC GAA ATT GCC TGG ATG GAC GGC AAT  1145
328     H   D   L   G   I   F   H   F   D   Q   I   A   A   W   G   A   G   E   I   A   W   M   D   G   N    352
1146   CTG AAG GGC TTT CCG GGG CGC GCC AGT GGC GAC CGC TGG GTG GTT CAG GCG AAG ATC ATC GTC GAG GAA GGG CTT  1220
353     L   K   G   F   P   G   R   A   S   G   D   R   W   V   V   Q   A   K   I   I   V   E   E   G   L    377
1221   GAC CGT TTC CTC GAA CGG GCG AAG ACA AAC AAT TAC TGA agacggaatactgatctgaagacttgaacgacaacctacagaggccgg  1307
378     D   R   F   L   E   R   A   K   T   N   N   Y   *                                                   390
1308   gacggcgtgacgtccgccacccggcacggaaagaggggaaaggaaaggcaaa ATG TCG ACG AAC GAC ACA GAC TGC CAG ACG AAA GCC  1394
1                                                     OrfA  M   S   T   N   D   T   D   C   Q   T   K   A    12
1395   CGC AAG CTG GGT GGT TGG CTC GGA GCC GCG ATT TTC TTT GCA CTT TGG CTT CTG GGT TCG GTG GGG ATC TGG ATG  1469
13      R   K   L   G   G   W   L   G   A   A   I   F   F   A   L   W   L   L   G   S   V   G   I   W   M    37
1470   GCG GCG ATC GCC GGG GTG ATC TCT GCG CTG ATC CTG ATC CTG ATC TGC TGG GCG ACC TGC GGT CTG CCC GAG  1544
38      A   A   I   A   G   V   I   S   A   L   I   L   A   R   L   I   C   W   A   T   C   G   L   P   E    62
1545   CCG AAA GCG CCG ATG TCG GCC GAG GAA GGC GCC GCC GCG CTG CGT CGG GCC GCC GTT CCG GCC GCG GCG ACG GCG  1619
63      P   K   A   P   M   S   A   E   E   G   A   A   L   R   R   A   V   P   A   A   A   P   T   A   A    87
1620   GTG AAA GTG GGC GAC GAG GCG CCG AAA CCG GCG CCG AAA GCC GCC GAG CCG GTC AAG GCC GCG CCC GTG GCG GCG CCG  1694
88      V   K   V   A   E   A   P   K   P   A   P   K   A   A   E   P   V   K   A   A   P   V   A   A   P    112
1695   GTC AAG ACG CCC GAG CCG GTG GCC GCG GTG GCT GAA GCG CCG AAA CCG GCG GCG AAG CCC GCC CCG GCG CCC GCG  1769
113     V   K   T   P   E   P   V   A   A   V   A   E   A   P   K   P   A   A   K   P   A   P   A   P   A    137
1770   CCG AAA GGC AAG TCG AAG CCC CGT CGC GGC TGA gacagacccgcccggcgcctggggagaggcgctgacgcggcagaaaccatcccga  1858
138     P   K   G   K   S   K   P   R   R   G   *                                                           148
1859   gccctgcggggctcgggatgtcccggaccagatccggaagcgccgggctgcggcctgacgagggagaccgtgtgtatcggacagccgatgccgcgcggc  1958
1959   cggacaccggccccgccgaaaccgaggggaatggacATG CTC AAC GAT CAC GAT CGC ATC TTC ACC AAC ATC TAC GGG ATG CAC  2042
1                                        NUOF  M   L   N   D   H   D   R   I   F   T   N   I   Y   G   M   H   16
2043   GAC CGC TCG CTG AAG GGC GCG ATG GCG CGT GGT CAC TGG GAT GGC ACG GCG GCG ATC CTT GGC AAT GGC CGC GAC  2117
17      D   R   S   L   K   G   A   M   A   R   G   H   W   D   G   T   A   A   I   L   G   N   G   R   D    41
2118   TGG ATC ATC GAG CAG GTC AAG GCC TCG GGG CTG CGC GGC CGC GGC GGT GCG GGC TTT CCG ACC GGG CTC AAG TGG  2192
42      W   I   I   E   Q   V   K   A   S   G   L   R   G   R   G   G   A   G   F   P   T   G   L   K   W    66
2193   TCC TTC ATG CCG AAG CAA TCC GAC GGG CGC CCC TCG TTT CTG GTG ATC AAC GCC GAT GAA TCC GAG CCG GGC ACC  2267
67      S   F   M   P   K   Q   S   D   G   R   P   S   F   L   V   I   N   A   D   E   S   E   P   G   T    91
2268   TGC AAG GAC CGC GAA ATC ATG CGC CAC GAT CCG CAT ACG CTG ATC GAA GGC GCG CTG ATC GCC GGT TTC GCC ATG  2342
92      C   K   D   R   E   I   M   R   H   D   P   H   T   L   I   E   G   A   L   I   A   G   F   A   M    116
2343   GGC GCG CGG GCG GCC TAT ATC TAT ATT CGC GGC GAA TAT GTC CGC GAA AAG GAG GCC CTG CAG GCC GCC ATC GAC  2417
117     G   A   R   A   A   Y   I   Y   I   R   G   E   Y   V   R   E   K   E   A   L   Q   A   A   I   D    141
2418   GAA TGC TAC GAG GCC GGG CTG ATC GGC AAG AAC GCC TGC CGC TCG GAC TAT GAT TTC GAC GTC TAT CTG CAC CAC  2492
142     E   C   Y   E   A   G   L   I   G   K   N   A   C   R   S   D   Y   D   F   D   V   Y   L   H   H    166
2493   GGT GCG GGC GCC TAT ATC TGC GGC GAG GAA ACC GCG CTT CTG GAA AGC CTT GAG GGC AAG AAG GGG ATG CCG CGG  2567
167     G   A   G   A   Y   I   C   G   E   E   T   A   L   L   E   S   L   E   G   K   K   G   M   P   R    191
2568   ATG AAA CCG CCG TTC CCG GCG GGC TCG GGG CTT TAT GGC TGC CCG ACG ACG GTG AAC AAC GTC GAA TCC ATC GCC  2642
192     M   K   P   P   F   P   A   G   S   G   L   Y   G   C   P   T   T   V   N   N   V   E   S   I   A    216
2643   GTC GTG CCG ACG ATC CTG CGC CGT GGT GGT GCC TGG TTC GCG GGC ATC GGG CGG CCG AAC AAC ACC GGC GTC AAG  2717
217     V   V   P   T   I   L   R   R   G   G   A   W   F   A   G   I   G   R   P   N   N   T   G   V   K    241
2718   CTG TTC GCG ATG TCG GGG CAC GTC AAC ACC CCC TGC GTC GTC GAA GAA GCC ATG TCG ATC AGC ATG AAG GAG CTG  2792
242     L   F   A   M   S   G   H   V   N   T   P   C   V   V   E   E   A   M   S   I   S   M   K   E   L    266
2793   ATC GAA AAG CAC GGC GGC GGC GTC CGC GGC TGG AAG AAC CTG AAA GCG GTG ATC CCC GGC GGC GCC TCC TGC  2867
267     I   E   K   H   G   G   G   V   R   G   W   K   N   L   K   A   V   I   P   G   G   A   S   C    291
2868   CCG GTC ATT CCG GCG CAT CTG TGC GAA GAC GCG ATC ATG GAT TAC GAC GGC ATG CGC GAA AAG CAG TCG TCC TTC  2942
292     P   V   I   P   A   H   L   C   E   D   A   I   M   D   Y   D   G   M   R   E   K   Q   S   S   F    316
2943   GGC ACC GCC TGT GTG ATC GAT CAG CAG ACC ATC AAG AAG GCG ATC GCC GGG CTG TCG AAA TTC TTC  3017
317     G   T   A   C   L   I   V   M   D   Q   Q   T   D   I   I   K   A   I   A   R   L   S   K   F   F    341
3018   AAG CAC GAG AGC TGC GGC CAG TGC ACG CCC TGC CGC GAG GGG ACC TCG TGG ATG TGG CGG GTG ATG GAG CGA CTG  3092
342     K   H   E   S   C   G   Q   C   T   P   C   R   E   G   T   S   W   M   W   R   V   M   E   R   L    366
3093   GTC ACC GGC GAG GCC GAG GTC GAG GAG ATC GAC ATG CTG TTT TCG GTG ACG AAA CAG GTC GAG GGC CAC ACG ATC  3167
367     V   T   G   E   A   E   V   E   E   I   D   M   L   F   S   V   T   K   Q   V   E   G   H   T   I    391
3168   TGC GCC CTG GGC GAT GCC GCC GCC TGG CCG ATC CAG GGG CTG ATC CGG CAT TAC CGC GAG GAG ATC GAG GAC CGG  3242
392     C   A   L   G   D   A   A   A   W   P   I   Q   G   L   I   R   H   Y   R   E   E   I   E   D   R    416
3243   ATC AAG GCG CGC AAG ACC GGC CGC ATG GGC GCG ATG GCG GCG GAG TGA tctgagatggcca  3303
417     I   K   A   R   K   T   G   R   M   G   A   M   A   A   E   *                   432
```
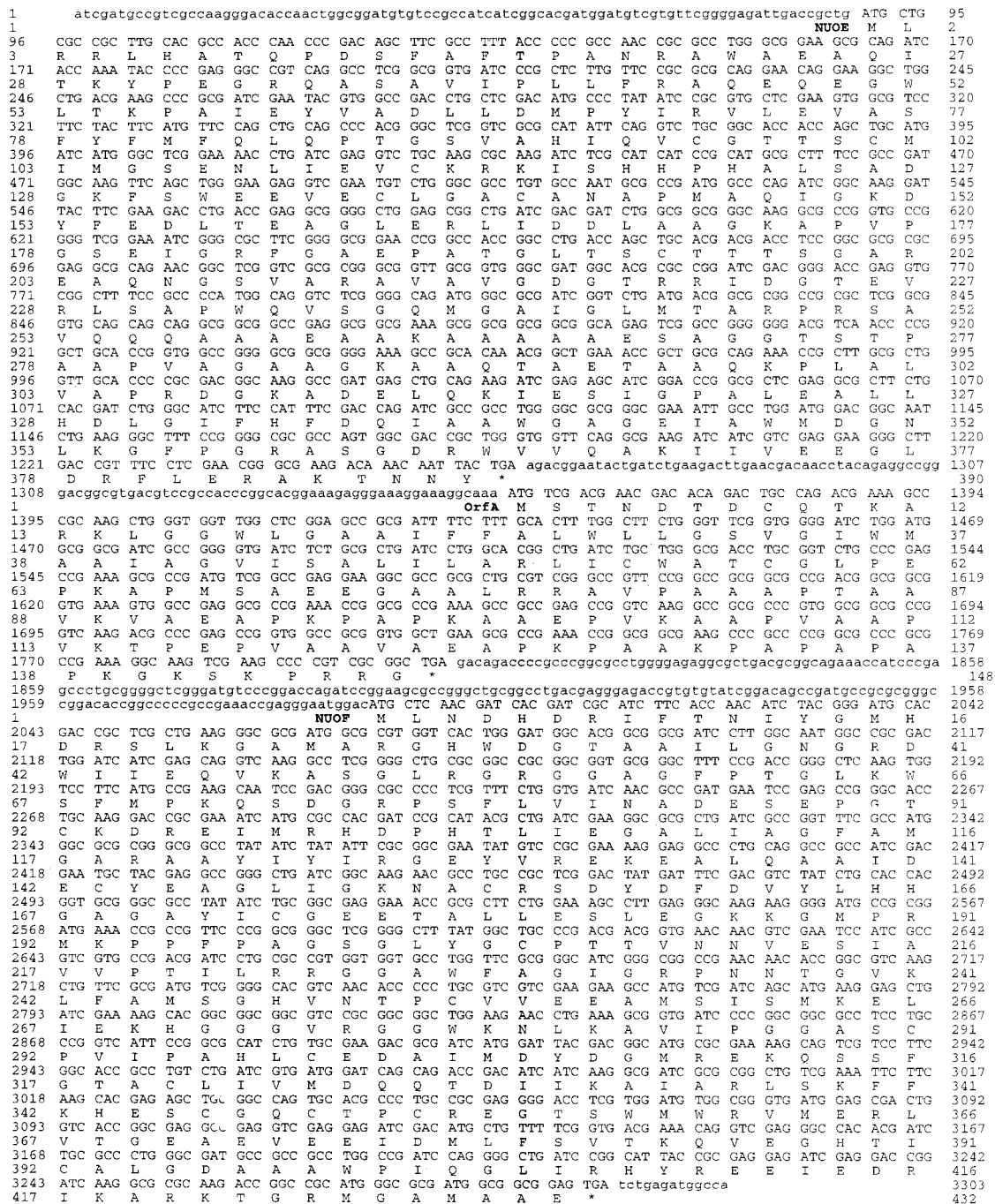
Fig. 1. Sequence of the *Cla*I-*Msc*I restriction fragment of lambda clones, selected from an *R. capsulatus* genomic library by screening with a cDNA coding for the bovine 51 kDa subunit as a probe. The derived amino acid sequences of the three open reading frames *nuoE*, *orfA* and *nuoF* are indicated below the DNA sequence. Termination codons are marked by an asterisk. The nucleotide sequence data reported in this figure will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under accession number Y10142.

onto the column and allowed to recirculate through the column over-night at 4°C. The column was then washed extensively in buffer A. Finally, the immuno-adsorbed components were released by an acidic washing step with a solution composed of 100 mM HCl-glycine pH 2.8, 0.4% Glucopon. Collected fractions were immediately neutralized and analyzed on SDS-PAGE. Proteins were sequenced after electro-phoretic transfer onto PVDF membranes as described [14].

The identification of the peptides generated by tryptic cleavage of NUOE was performed by mass spectrometry as follows. The silver stained NUOE spot was excised from the polyacrylamide gel and in-gel digested with trypsin after iodoacetamide derivatization according to published procedures [15]. The resultant peptide mixture was em-bedded into a matrix of α-cyano-4-hydroxy-trans-cinnamic acid and analyzed by MALDI mass spectrometry. The mass spectra were ob-tained on a Bruker BIFLEX mass spectrometer. For spectrum cali-bration, trypsin autolytic fragments were taken as standards.

3. Results and discussion

Lambda clones, containing a DNA fragment encompassing the R. capsulatus gene coding for the 51 kDa catalytic subunit of the respiratory complex I, were recovered by screening and R. capsulatus DNA genomic library using a cDNA clone cod-ing for the bovine mitochondrial 51 kDa subunit as a probe. Restriction digest fragments of the positive clones were cloned and sequenced. A ClaI-MscI fragment 3.3 kb long (Fig. 1) was proved to contain 3 putative open reading frames. The first open reading frame from base 90 to base 1259 encodes a 41 kDa protein which is related to the bovine 24 kDa subunit (Fig. 2). As a consequence this gene was called nuoE following the nomenclature of the E. coli nuo genes [10]. The second open reading frame from base 1359 to 1802 was called orfA and encodes a potential polypeptide with a molecular mass of 14.8 kDa. This putatively expressed polypeptide has no equiv-alent with either of the known subunits of the mitochondrial or bacterial complexes I. The last open reading frame from base 1995 to base 3290 encodes a 47 kDa protein which is equivalent to the bovine 51 kDa subunit. This gene was called nuoF. The overall organization of this R. capsulatus DNA fragment with the nuoE gene located upstream of the nuoF gene is similar to what is found in the E. coli nuo operon, except that in R. capsulatus a supernumerary open reading

frame (orfA) is located between nuoE and nuoF. In the P. denitrificans operon, Yagi's group has discovered two open reading frames called urf1 and urf2 located between the nqo2 and nqo1 genes (nqo2 and nqo1 genes are the equivalents of the R. capsulatus nuoE and nuoF genes respectively) [9]. Neither of these two urfs looks similar to the R. capsulatus orfA. Analysis of our lambda clones showed that in R. cap-sulatus the nuoE and nuoF genes are located in the close vi-cinity and upstream of the nuoH-nuoN gene cluster previously sequenced (not shown) [11,16].

A closer look at the primary structures of the NUOE and NUOF proteins has given additional information. The R. capsulatus NUOF protein is 38.5% identical to the E. coli NUOF subunit, about 67% identical to the bovine 51 kDa subunit and up to 85% identical to the P. denitrificans NQO1 subunit. Four cysteine residues which are believed to be involved in the binding of a 4Fe-4S cluster (C346, C349, C352 and C392, R. capsulatus numbering) are strictly con-served in all the subunits. Four cysteine residues are also in-variant in the NUOE equivalent subunits (Fig. 2). All the NUOE equivalent subunits are highly similar except for their C terminal parts which are heterogeneous in size (Fig. 2). The E. coli NUOE subunit is the shortest one with 166 amino acids and a molecular mass of about 19 kDa. The human and bovine subunits, as mitochondrially imported and proc-essed forms, are 217 amino acids and close to 24 kDa, and the P. denitrificans with 239 amino acids is 26 kDa [9,10,17]. Ac-cording to the DNA sequence, the R. capsulatus NUOE sub-unit would have 389 amino acids and a molecular mass of 41 264 Da. This subunit looks very atypical and is nearly twice as big as the corresponding mammalian mitochondrial sub-unit. In comparison to the P. denitrificans subunit, the R. cap-sulatus NUOE subunit appeared to be extended at the C terminus by 150 amino acids. This predicted extension is ala-nine rich, without specific motifs, and contains only one addi-tional cysteine residue. It is not highly hydrophobic and is not predicted to form a transmembranous helix. To rule out the possibility of any DNA sequencing artifact, we paid careful attention to accurate sequencing of various subcloned frag-ments of the nuoE gene as well as a number of PCR fragments
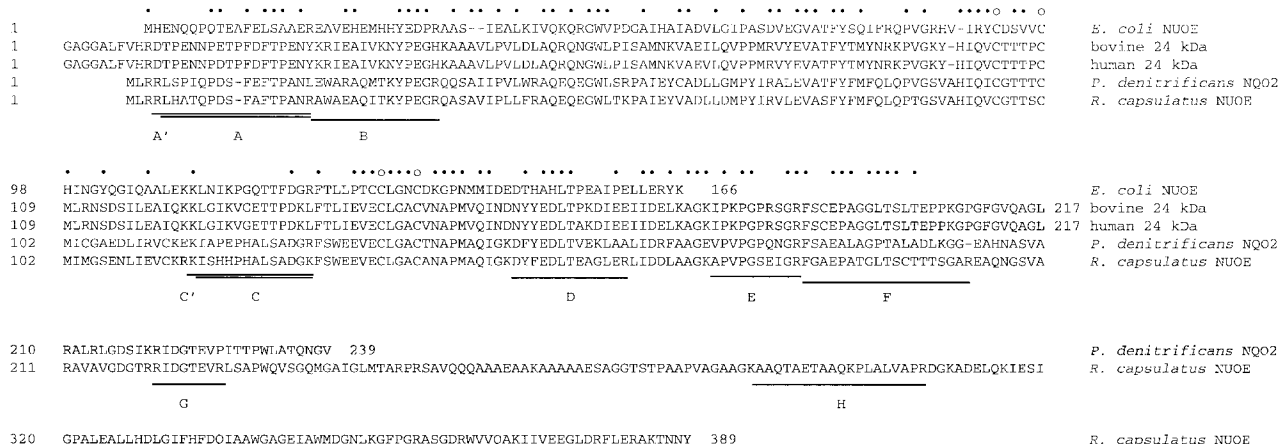
Fig. 2. Comparison of the polypeptide sequences of the R. capsulatus NUOE subunit, the E. coli NUOE subunit, the bovine and human 24 kDa subunits (as processed mitochondrially imported subunits) and the P. denitrificans NQO2 subunit. Identical residues between the R. capsu-latus NUOE and the human 24 kDa subunits are marked by dots, invariant cysteine residues are indicated by open circles. The tryptic peptides of the R. capsulatus NUOE subunit identified by MALDI mass spectrometry (see Fig. 4) are underlined and labelled A to I and A' to C'. Alignment was computerized using the Clustal method with the PAM 250 weight table.
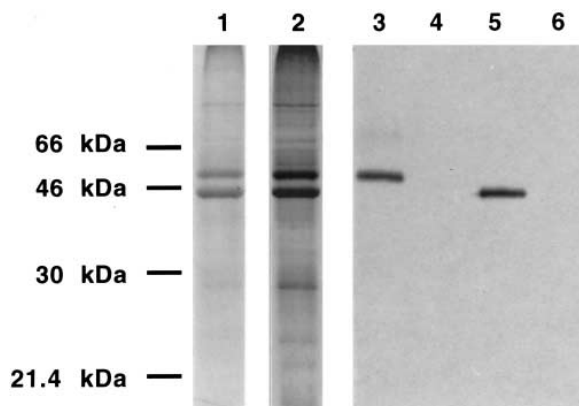
Fig. 3. SDS polyacrylamide gel and Western blot analysis of the fractions eluted from the anti-PNE affinity column. Proteins recovered from the affinity column by acidic elution were precipitated overnight with 5 volumes of ethanol at −20°C (lanes 1, 2, 3 and 5). Membrane lysates collected in the flow through of the column during sample loading were also run as controls after being ethanol precipitated (lanes 4 and 6). Samples were run through a 12% polyacrylamide gel then stained with Coomassie (lane 1) or silver (lane 2). After electrophoretic transfer to PVDF membranes, proteins were identified using an anti-NUOF antibody (lanes 3 and 4) or an anti-PNE antibody (lanes 5 and 6).

overlapping this gene. No error was found in our consensus sequence. Accordingly, we set up immunoaffinity experiments to check the size of the expressed subunit in the R. capsulatus membranes so as to definitively ascertain the existence of an elongated C terminus in the R. capsulatus NUOE subunit.

Antibodies were raised against a synthetic peptide (called PNE) corresponding to the deduced N terminus sequence of the NUOE subunit. Membrane samples from the R. capsulatus B10 strain were analyzed by 2D electrophoresis in polyacrylamide gel. The anti-PNE antibodies specifically revealed one polypeptide with an apparent molecular mass of 42 kDa and a pI value close to 5.5 (not shown). These features are very close to the theoretical values of 41 264 Da and pI 5.42 deduced from the sequence of the gene. The selectivity of our antibody allowed us to set up experiments with the aim to immuno-purify the whole complex I or some fraction of it from R. capsulatus membrane fractions. Membranes were diluted in a 2% Glucopon solution (see Section 2) and soluble material was recovered after centrifugation. This soluble fraction was loaded onto a column of anti-PNE antibodies immobilized on a protein A-crosslinked Sepharose. After washing, the eluted fractions (see Section 2) were analyzed on SDS-PAGE. These fractions contained 2 major proteins in a grossly stoichiometric amount as revealed by Coomassie blue staining (Fig. 3, lane 1). This subcomplex fraction appeared at least 95% pure as shown by Coomassie and intense silver staining (Fig. 3, lanes 1 and 2). Apparent molecular masses of 49 and 43 kDa were estimated for these two proteins. The two proteins were specifically detected by an anti-NUOF subunit antibody or by the anti-PNE antibody respectively (Fig. 3). The identity of these two proteins was further confirmed by sequencing of their N termini. The N terminus of the biggest subunit was shown to be MLNDHDRIFT and that of the smaller one MLRRLHATQP. These N termini perfectly matched those deduced from the sequenced genes coding for the NUOF and NUOE subunits respectively

(Fig. 1). This confirmed that the two newly sequenced nuoE and nuoF genes are truly expressed in R. capsulatus and that the NUOE and NUOF subunits, purified as a soluble heterodimer after detergent extraction of the membranes, are in close contact in complex I. We also checked on an aliquot of the particulate material from the column if the NUOE-NUOF subcomplex might have some catalytic activity. Indeed, some deamino-NADH dependent nitroblue tetrazolium reductase activity was present in the immobilized protein fraction on the column after the extensive washes.

To ascertain that the mass of the NUOE subunit was related to a C terminus extension in comparison to other species NUOE subunits, we analyzed by MALDI mass spectrometry the peptide mixture obtained after tryptic digestion of the NUOE spot in the polyacrylamide gel. Eleven different peptides generated by trypsin cleavage of the NUOE subunit were unambiguously identified (Fig. 4). Among these peptides, two (peptide H from amino acid 288 to 306 and peptide I from amino acid 371 to 379) are part of the C terminus extension which has no counterpart in the equivalent bacterial or mitochondrial subunits (Figs. 2 and 4). This analysis ascertains that the amino acid sequence deduced from the DNA sequence of the nuoE gene is accurate and that the R. capsulatus NUOE subunit is significantly bigger than the other equivalent subunits known so far.

In chloroplasts, three subunits of the NADH dehydrogenase domain of complex I (75 kDa, 51 kDa and 24 kDa) are apparently missing. This observation was taken by Friedrich et al. as an indication that in chloroplasts 'complex I' may work in a different way to the mitochondrial complex I [18]. Chloroplast complex I was hypothesized by these authors to be a NADPH or a ferredoxin:plastoquinone oxidoreductase. The discovery that in R. capsulatus NUOE and NUOF, which are the equivalents of the 24 kDa and 51 kDa subunits, are expressed is a clear indication that the R. capsulatus complex I is more of the mitochondrial type than of the chloroplast type.

Friedrich et al. have shown that stripping the E. coli cytoplasmic membranes under acidic conditions with a buffer containing the technical alkylglucoside APG 225 (a Glucopon equivalent manufactured by Henkel, Düsseldorf, Germany) led to the isolation of a whole active complex I [19]. Unlike this work, many different attempts at solubilizing the whole R. capsulatus complex I in our laboratory or others [20] have thus far proved unsuccessful probably because of the great instability of this complex I. The experiments we report here clearly show that in the case of R. capsulatus, Glucopon treatment of the membranes led to the isolation of a subcomplex composed of the NUOE and NUOF subunits only. This subcomplex was stripped from the membrane whatever the pH of the stripping solution in the range of 6–8. Indeed, our purification protocol yielded a subcomplex of complex I very similar to the mitochondrial FP fragment. The mitochondrial FP fragment is composed of 3 different subunits, 51 kDa, 24 kDa and 9 kDa. An equivalent of the 9 kDa subunit, the role of which is unclear in the mitochondrial complex I, is absent in our preparation. None of the genes or other ORFs of the bacterial NADH-ubiquinone oxidoreductase operons analyzed so far encode a protein similar to the mammalian 9 kDa ([9,10] and unpublished data for R. capsulatus). Furthermore, no significant homology was observed between the C terminal extension of R. capsulatus NUOE subunit and the bovine 9 kDa. Consequently, the possibility for the R. capsu-
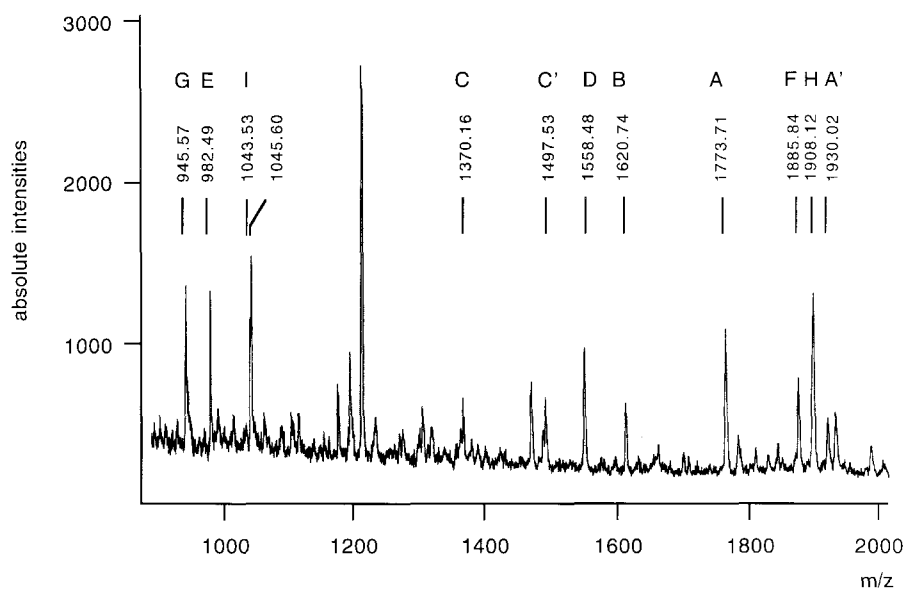
Peptide	Position	Sequence	Theoretical MH+
G	221-228	RIDGTEVR	945.51
E	174-183	APVPGSEIGR	982.53
I	371-379	IIVEEGLDR	1043.57
C	117-129	ISHHPHALSADGK	1369.70
C'	116-129	KISHHPHALSADGK	1497.79
D	152-164	DYFEDLTEAGLER	1557.71
B	21-34	AWAEAQITKYPEGR	1619.82
A	5-20	LHATQPDSFAFTPANR	1772.87
F	184-202	FGAEPATGLTSCTTTSGAR	1884.88 (a)
H	288-306	AAQTAETAAQKPLALVAPR	1907.07
A'	4-20	RLHATQPDSFAFTPANR	1928.97

Fig. 4. MALDI mass spectrum of a tryptic digest of the NUOE subunit. After in-gel tryptic digestion of the NUOE band, the eluted peptide mixture was analyzed as described in Section 2. The identified tryptic peptides are labelled A to I, and A' and C', their positions and sequences are given in the table. Peptides A', C' and G correspond to partially digested peptides. The peptide at m/z 1045.60 is a trypsin autolytic fragment which partially overlaps the signal of peptide I. These two peaks are clearly distinguishable when this doublet peak is magnified (not shown). The major peak of the spectrum was unidentified. The theoretical masses of the protonated peptides are given in the table. They correspond to the monoisotopic masses for peptides G, E and I and to average masses for peptides A–D, C', F and H. Positions of the peptides A to I, A' and C' are underlined in Fig. 2. (a) Mass of the Cys-195 carboxyamidomethylated peptide F.

latus NUOE subunit to behave as a fusion protein that would be the equivalent of both the bovine 24 kDa and 9 kDa seems to be ruled out. In the case of *P. denitrificans* attempts to purify complex I after detergent treatment of the membranes have led to the partial purification of a heterodimer with NADH-ubiquinone oxidoreductase activity [21]. The two subunits had molecular masses of 48 and 25 kDa. An antibody raised against this heterodimer was proved to crossreact with the bovine 51 kDa subunit of the bovine mitochondrial complex I suggesting that the two-subunit NADH-ubiquinone oxidoreductase component purified from the *P. denitrificans* membranes might also be an FP-like subcomplex I [22]. Furthermore, it has been recently reported by Yagi's group that the *P. denitrificans* NQO1 and NQO2 subunits tend to associate in a dimeric aggregate when co-expressed in *E. coli* [23]. The above data as well as our immuno-purification of an FP-like fragment from the *R. capsulatus* membrane are clear indications that the NUOE and NUOF subunits interact closely and strongly in the bacterial complex I.

The *R. capsulatus* NUOE subunit has a very peculiar feature because of a very long C terminal alanine rich extension which is not found in the equivalent subunits from other species. Because this extension is predicted not to contain any hydrophobic helix it is probably not involved in the membrane anchorage of complex I. At the present time, we do not know if this extended C terminus is related to a specific structural component in the *R. capsulatus* complex I or whether it is associated with the fact that *R. capsulatus* is a photosynthetic bacterium. Mutagenesis studies are under way to check the role, if any, of this C terminal extension.

References

[1] I.M. Fearnley, J.E. Walker, Biochim. Biophys. Acta 1140 (1992) 105–134.
[2] T. Yagi, Biochim. Biophys. Acta 1141 (1993) 1–17.
[3] Y. Hatefi, Annu. Rev. Biochem. 54 (1985) 1015–1069.

[4] C.I. Ragan, Curr. Topics Bioenerg. 15 (1987) 1–36.

[5] Y.M. Galante, Y. Hatefi, Arch. Biochem. Biophys. 192 (1979) 559–568.

[6] C.I. Ragan, Y.M. Galante, Y. Hatefi, T. Onishi, Biochemistry 21 (1982) 590–594.

[7] S. Chen, R.J. Guillory, J. Biol. Chem. 256 (1981) 8318–8323.

[8] T. Yagi, T.M. Dinh, Biochemistry 29 (1990) 5515–5520.

[9] T. Yagi, T. Yano, A. Matsuno-Yagi, J. Bioenerg. Biomemb. 25 (1993) 339–345.

[10] U. Weidner, S. Geier, A. Ptock, T. Friedrich, H. Leif, H. Weiss, J. Mol. Biol. 233 (1993) 109–122.

[11] A. Dupuis, FEBS Lett. 301 (1992) 215–218.

[12] P.F. Weaver, J.D. Wall, H. Gest, Arch. Microbiol. 105 (1975) 207–216.

[13] A.M.R. Haines, J.M. Cooper, J.A. Morgan-Hughes, J.B. Clark, A.H.V. Schapira, Prot. Express. Purif. 3 (1992) 223–227.

[14] C. Adessi, A. Chapel, M. Vinçon, T. Rabilloud, G. Klein, M. Satre, J. Garin, J. Cell. Sci. 108 (1995) 3331–3337.

[15] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Anal. Chem. 68 (1996) 850–858.

[16] A. Dupuis, A. Peinnequin, M. Chevallet, J. Lunardi, E. Darrouzet, B. Pierrard, V. Procaccio, J.-P. Issartel, Gene 167 (1995) 99–104.

[17] S.J. Pilkington, J.E. Walker, Biochemistry 28 (1989) 3257–3264.

[18] T. Friedrich, K. Steinmüller, H. Weiss, FEBS Lett. 367 (1995) 107–111.

[19] H. Leif, V.D. Sled, T. Onishi, H. Weiss, T. Friedrich, Eur. J. Biochem. 230 (1995) 538–548.

[20] B.C. Berks, S.J. Ferguson, Biochem. Soc. Trans. 19 (1991) 581–588.

[21] C.L. George, S.J. Ferguson, Eur. J. Biochem. 143 (1984) 567–573.

[22] C.L. George, S.J. Ferguson, M.W.J. Cleeter, C.I. Ragan, FEBS Lett. 198 (1986) 135–139.

[23] T. Yano, V.D. Sled, T. Ohnishi, T. Yagi, J. Biol. Chem. 271 (1996) 5907–5913.