Volume 301, number 2, 215–218 FEBS 10948 April 1992

Identification of two genes of *Rhodobacter capsulatus* coding for proteins homologous to the ND1 and 23 kDa subunits of the mitochondrial Complex I

Alain Dupuis

Laboratoire de Biochimie URA CNRS 1130, Département de Biologie Moléculaire et Structurale, Centre d'Etudes Nucléaires de Grenoble, 38041 Grenoble Cedex, France

Received 9 March 1992

A region of the genome of *Rhodobacter capsulatus* has been sequenced and shown to encode proteins homologous to the ND1 subunit and the 23 kDa subunit of the mitochondrial NADH:CoQ oxidoreductase (Complex I). The association of these two open reading frames in the *R. capsulatus* genome parallels the organisation of the chloroplast genome. It suggests that genes encoding subunits of the NADH:CoQ oxidoreductase must be clustered in the genome of *R. capsulatus*.

Rhodobacter capsulatus; NADH dehydrogenase; NADH:Q oxidoreductase; Complex I; Mitochondrion

1. INTRODUCTION

The NADH:CoQ oxidoreductase (Complex I) is the first complex of the mitochondrial respiratory chain. It catalyses electron transfer from NADH to ubiquinone Q10 in a process coupled with an active efflux of protons through the mitochondrial inner membrane. This enzyme is made-up of at least 35 subunits and contains one FMN cofactor and 4–6 iron–sulfur clusters (for a review see [1]). Seven of the subunits of the Complex I are encoded by the mitochondrial genome [2,3] while the remaining are encoded by the nucleus. Due to its high complexity, Complex I is the least understood of the mitochondrial respiratory complexes. Similar to the other respiratory complexes, the study of bacterial NADH:CoQ oxidoreductases may no doubt allow decisive progress in the understanding of the mitochondrial Complex I for the following reasons: (i) bacterial respiratory complexes are usually simpler than their mitochondrial counterparts [4,5]; (ii) bacterial genome organisation may allow faster progress in the cloning of the genes encoding the different subunits of the putative Complex I; (iii) bacteria allow for genetic studies hardly accessible with the mitochondrial enzyme. Actually, bacteria present two different types of NADH:CoQ oxidoreductases (for reviews see [6,7]). The type-II NADH:CoQ oxidoreductases show little relation with the mitochondrial Complex I; they are small dimeric enzymes with an associated FAD cofactor and no iron–sulfur clusters. On the other hand, the type-I NADH:CoQ oxidoreductases display a great similarity with the mitochondrial complex I [6]. The type-I NADH:CoQ oxidoreductase of non-sulfur α-purple bacteria such as *R. capsulatus* and *Paracoccus denitrificans* is sensitive to rotenone [6,7] and shows EPR signals very similar to those associated with the mitochondrial Complex I [8–10]. This enzyme has been purified from *P. denitrificans* [11]. The genes encoding two of its subunits were recently cloned and proved to be homologous to the genes encoding the 24 kDa and the NADH-binding subunits of the mitochondrial Complex I [12,13]. In contrast, the type-I NADH:CoQ oxidoreductase of *R. capsulatus* appeared resistant to purification and biochemical characterisation [7]. The present paper deals with the cloning and characterisation of two genes of *R. capsulatus* which encode proteins highly homologous to the ND1 subunit and the 23 kDa subunit of the mitochondrial Complex I.

Correspondence address: A. Dupuis, Laboratoire de Biochimie URA CNRS 1130, DBMS, Centre d'Etudes Nucléaires de Grenoble, BP 85X, 38041 Grenoble Cedex, France. Fax: (33) (76) 88 5185.

2. MATERIALS AND METHODS

Genomic DNA was isolated from exponential cultures of *R. capsulatus* strain B10 according to Marmur [14]. DNA fragments were cloned in phage M13mp18 or in the bluescript phagemids KS$^+$ or SK$^+$ (Stratagene, USA). Polymerase chain reaction (PCR), Southern blotting and other genetic manipulations were done as described previously [15,16]. Generation of unidirectional deletions in the cloned DNA and production of single strand DNA phagemids were realised according to the procedure described by Stratagene. DNA sequencing was performed using the Pharmacia T7 DNA polymerase kit. Sequences were compiled, analysed and compared to the protein SWISS-PROT and PIR (NBRF) databases using the computer programs package DNASTAR (DNASTAR, UK).

3. RESULTS AND DISCUSSION

We have observed that the 23 kDa subunit of bovine Complex I displays significant similarities both with the chloroplast frxB proteins and with the bacterial 8Fe–8S ferredoxins [15]. These observations allowed us to design degenerate oligonucleotides for the use as primers in a PCR experiment with *R. capsulatus* genomic DNA template. The forward degenerate primer corresponded to the protein consensus sequence $_{47}$TI(H/Q/N)YP(Y/F)EK$_{54}$ deduced from comparison of the mitochondrial 23 kDa with the chloroplast frxB proteins (numbering corresponds to the mitochondrial 23 kDa protein). For the design of the reverse primer, we chose the protein sequence $_{126}$CPVDA(I/P)$_{131}$ of the bovine 23 kDa subunit which shows a remarkable degree of conservation among bacterial 8Fe–8S ferredoxins. The product of the PCR was analysed by Southern blotting with the degenerate probe PH23 designed after the conserved protein sequence $_{77}$CIAC(E/K)(L/V)$_{82}$ corresponding to the first iron–sulfur insertion site of the bovine 23 kDa subunit. The PCR product yielded a strong hybridisation signal and was cloned into M13mp18 and sequenced. In one of the six reading frames, the cloned DNA fragment was shown to potentially encode for a peptidic sequence 81% similar to the central part of the 23 kDa subunit of the mitochondrial Complex I. This PCR product was used as a probe to analyse different restriction digests of genomic DNA of *R. capsulatus* (Fig. 1). The 1.6 kb fragment obtained by double digestion with the enzymes *Bam*HI–*Hin*dIII and the contiguous 300 bp *Hin*dIII fragment which hybridised with the probe (see Fig. 1) were cloned in bluescript and M13 vectors respectively. The initial PCR product overlapped both this short fragment and the *Bam*HI–*Hin*dIII sequence thus giving a continuous DNA fragment 1849 bases long (Fig. 2).

Taking into account the codon usage bias characteristic of Rhodospirillaceæ [17] the sequence of the cloned DNA was searched in both directions for open reading frames longer than 100 bases. Three potential coding frames (orf) were identified which displayed ribosome binding sites (underlined in the Fig. 2) upstream to their start codon [18]. As expected, the protein encoded by the open reading frame from base 1338 to base 1826 proved highly homologous to the bovine 23 kDa subunit as well as the chloroplast frxB proteins (see Fig. 3). In view of the relationships with the 23 kDa subunit of mitochondrial Complex I and to be consistent with the nomenclature previously used for chloroplast *ndh* genes [15] we propose to call this gene *ndhI*. The corresponding NDH-I protein is 163 amino acids long (M.W. 18,848 Da). It is a ferredoxin-like protein characterised by the presence of two sequence motifs of the form CxxCxxCxxxCP (Figure 3) typical of 4Fe–4S cluster insertions sites. Thus this protein and its mitochondrial counterpart must contain two of the 4–6 iron–sulfur

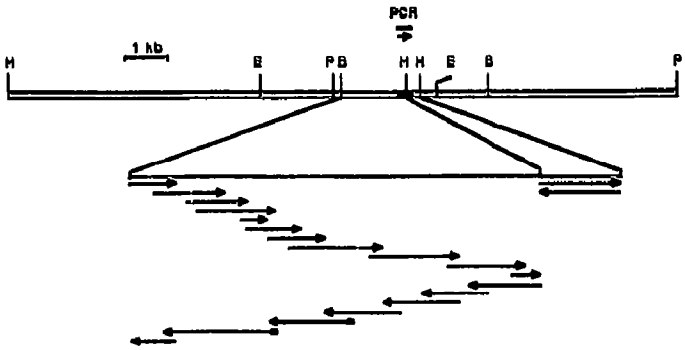

Fig. 1. Strategy for the cloning and sequencing of the *ndhA* and *ndhI* open reading frames. The initial PCR product potentially coding for a fragment of protein homologous to the mitochondrial 23 kDa subunit was used as a probe to analyse the restriction digests of *R. capsulatus* genome. This allowed us to design a restriction map of the region of the genome surrounding the probe. We cloned the 1.6 kb *Bam*HI–*Hin*dIII and the 0.3 kb *Hin*dIII fragments in Bluescript and M13 vectors, respectively (bold lines). The 1.6 kb fragment cloned in bluescript was sequentially deleted using the ExonucleaseIII-Mung Bean DNase deletion procedure. This allowed for the sequencing of quite the totality of the insert on both strands (→) using the universal M13 and T7 sequencing primers. One gap in the reverse strand sequence was sequenced using two specific primers. They are indicated by the two boxed arrows (▬→). The letters on the restriction map correspond to the following enzymes: B, *Bam*HI, E, *Eco*RI, H, *Hin*dIII and P, *Pst*I.

clusters identified in the mitochondrial and bacterial NADH:CoQ oxidoreductases.

The second open reading frame (from base 1057 to base 1335) was termed *orf* 1 (Fig. 2). It potentially codes for a 93 amino-acid long hydrophilic protein (M.W. 10,067 Da). The protein encoded by this orf displays no extended homology with any known protein of the SWISS-PROT and NBRF PIR databases or with known proteins of mitochondrial Complex I of similar molecular weight.

On the other hand, the third open reading frame (from base 17 to base 1051) corresponds to a 345 amino-acids long protein (M.W. 37,855 Da) homologous to the mitochondrial ND1 subunit or to the chloroplast NDH-A protein (see Fig. 4). Thus, this open reading frame should be named *ndhA*, and the corresponding peptide should be called NDH-A. In mitochondria this protein is located in the hydrophobic subfraction of the Complex I [1]. It is embedded in the mitochondrial inner membrane and bears the binding site(s) for ubiquinone and rotenoid inhibitors [21].

The seven mitochondrial genss (ND1 to ND6 plus ND4L) and three nuclear genes coding for the 49 kDa, the 30 kDa and the 23 kDa subunits of the mitochondrial Complex I display similarities with ten open reading frames (*ndhA–J*) identified in the genomes of chloroplast and blue algae [16,17,19,22–25]. These observations have led to the proposal that the putative chloroplast proteins encoded by these *ndh* genes are subunits of a NADH- or NADPH-plastoquinone reductase

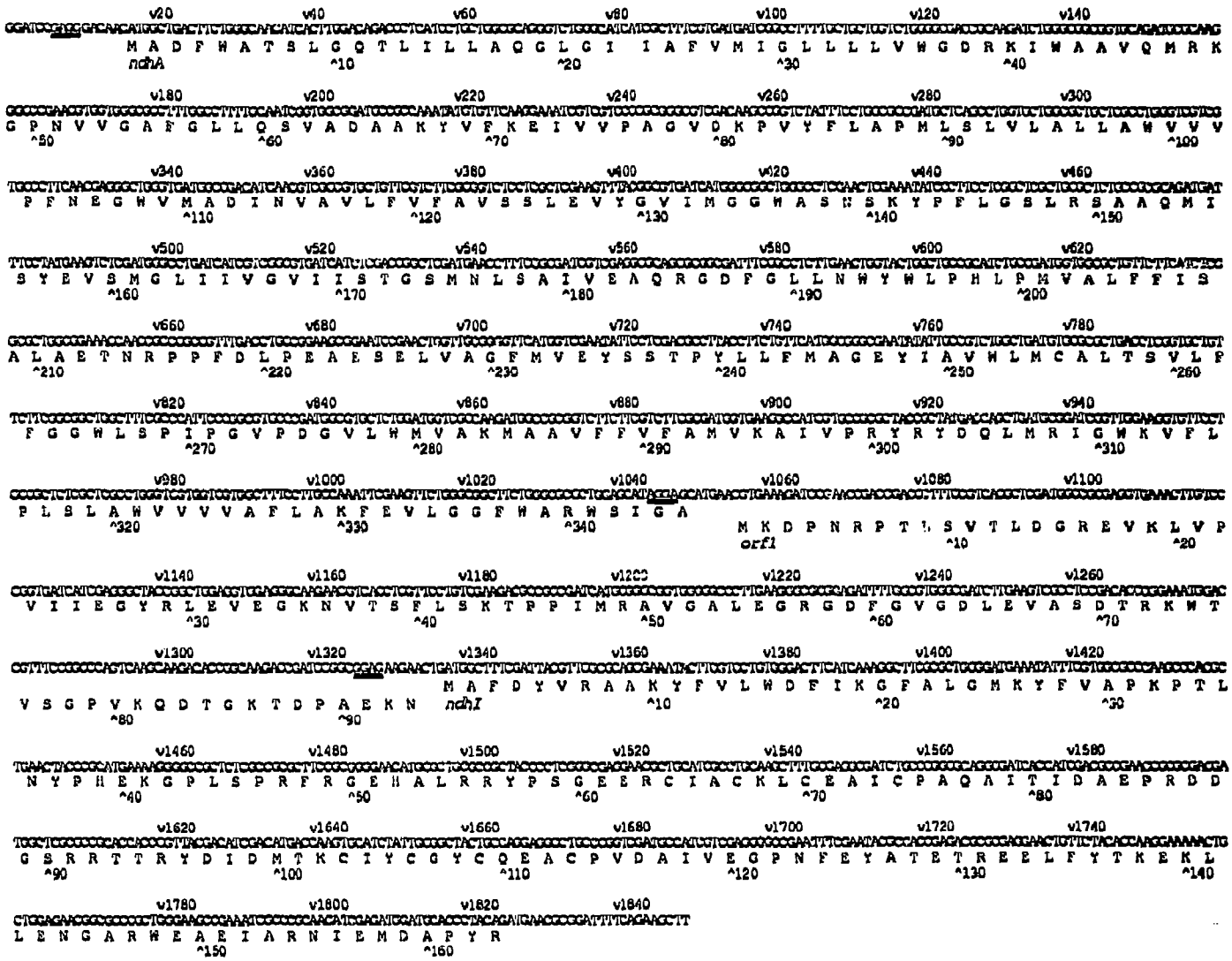

Fig. 2. Compiled sequence of the characterised DNA fragment and of its potential gene products. The DNA sequence of the 1.85 kb fragment was searched on both strands for open reading frames longer than 100 bases with codon usage bias characteristic of Rhodospirillaceæ [17]. Three open reading frames were identified. The protein sequence of their encoded products is given under the DNA sequence. The potential ribosome binding site associated with each open reading frame is underlined [18].

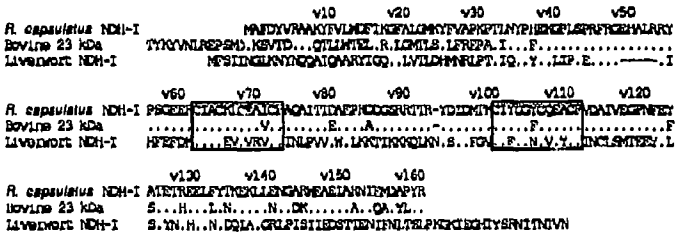

Fig. 3. Alignment of the protein sequences of the bacterial NDH-I with the mitochondrial 23 kDa subunit and chloroplast NDH-I proteins. The protein sequence of the NDH-I protein of *R. capsulatus* was aligned with the protein sequences of the 23 kDa subunit from bovine mitochondria [16] and the NDH-I protein from liverwort chloroplast [19]. The dots represent conserved residues. This representation was chosen to highlight the high degree of conservation between the mitochondrial and bacterial proteins. The boxed sequences correspond to the potential sites of insertion of iron-sulfur clusters.

closely related to the mitochondrial Complex I [23]. In the chloroplast genome, clustering of *ndh* genes seems reminiscent of an ancestral bacterial operon. In this regard the association of *ndhA* and *ndhI* genes in the genome of *R. capsulatus* parallels the organisation of the chloroplast genome. Cloning of the regions flanking the *ndhI* and *ndhA* genes of *R. capsulatus* is under progress. It will allow us to see if the similitude of organisation extends to the other *ndh* genes.

Extensive experimental evidence points to the similarity between the mitochondrial Complex I and the Type I NADH:CoQ oxidoreductase of *R. capsulatus* (reviewed in [7]). However, due to its lability, this enzyme has not been isolated yet and little is known of its molecular characteristics. The present study marks a decisive step in the molecular characterisation of the Type I

```
                                    v1        v10       v20       v30       v40       v50       v60
R. capsulatus NDH-A                 MADFWATSLGQTLILLAQGLGIIAFVMIGLLLLVWGDRKIWAAVQMRKGPNVVGAFGLLQSVADA
Human mitochondrial ND1                                 MPM.NL.LL.VPIL.AMAF.MLTE...LGYM.L........PY....PF...
Liverwort chloroplast NDH-A MISNINLEDKFFSFFFTLGFSKEFFNF.WIIFSI.ILMLG.T..V.V...LE...S..I.Q.I..EYA.PL.II.AL..G

                                  v70       v80       v90       v100      v110      v120      v130      v140
R. capsulatus NDHA                  AKYVFKEIVVPAGVDKPVYFLAPMLSLVLALLAWVVVPFNEGWVMADINVAVLFVFAVSSLEVYGVIMGGWASNSKYPFL
Human mitochondrial ND1             M.LFT..PLK..TSTITL.IT..T.A.TI...L.TPL.MPNPL.NL--.LGL..IL.T...A..SILWS......N.ALI
Liverwort chloroplast NDH-A         I.LFL..DIV..QG.VWLFNIG.I.V.IPVF.SYLYNVILANFSI-----G.F.WI....VVPL.LL.A.YG..N..S..

                                  v150      v160      v170      v180      v190      v200      v210      v
R. capsulatus NDH-A                 GSLRSA-----AQMISYEVSMGLIIVGVIISTGSMNLSAIVEAQRGDFGLLNWYWLPHLFMVALFFISALAETNRPPFDL
Human mitochondrial ND1             .A..AV-----..T.....TLAI.LLSTLLMS..F...TLITT.EHL---.-.LL..SW.LAMMW...T......T....
Liverwort chloroplast NDH-A         .G..AVIPFEA..S....IPLA.SVLSIALLSN.LSTVD.....SKYGF.SWNL.RQPIGFI-V...AS...CE.L....

                                  v230      v240      v250      v260      v270
R. capsulatus NDH-A                 PEAESELVAGFMVEYSSTPYLLFMAGEYIAVWLMCALTSVLFFGGWLSPIP----------------------GVPDGVLW
Human mitochondrial ND1             AEG.....S..NI..AAG.FAL.F....TNIIM.NT..TTI.L.TTYDALS----------------------PELYTTY
Liverwort chloroplast NDH-A         ....E.....YQT...GMKFAF.YLAS.LNLLVSSLFVTI.YL...HFS..FFSLFKNFEWNLMSNGISEVISIII.IVI

                                  v280      v290      v300      v310      v320      v330      v340
R. capsulatus NDH-A                 MVAKMAAVFFVFAMVKAIVPRYRYDQLMRIGWKVFLPLSLAWVVVVAFLAKFEVLGGFWARWSIGA
Human mitochondrial ND1             F.T.TLLLTSL.LWIRTAY..F......HLL..N....T..LLMWYVSMPITISSIPPQT
Liverwort chloroplast NDH-A         TLV.SYLFL.ISI.TRWTL..I.I...LNL...FL..IA.GNLLLTTSFQL.LL
```

Fig. 4. Alignment of the protein sequences of the bacterial NDH-A with the mitochondrial ND1 subunit and the chloroplast NDH-A protein. The protein sequence of the NDH-A protein of *R. capsulatus* was aligned with the protein sequences of the ND1 subunits from human mitochondria [20] and the NDH-A protein from liverwort chloroplast [19]. The dots represent conserved residues.

NADH:CoQ oxidoreductase of *R. capsulatus*. It will allow the use of the powerful genetic tools developed in *R. capsulatus* for the study of this essential yet poorly understood enzyme.

Acknowledgements: I am very grateful to Drs J.E. Walker, J.-P. Issartel and J. Lunardi for critical advice and careful reading of this manuscript. This work was realised in the laboratory of Pr. Vignais. I am very grateful for his support. The sequence data reported in this paper have been submitted to EMBL Data library under the submission number Z11611.

REFERENCES

[1] Weiss, H., Friedrich, T., Hofhaus, G. and Preis, D. (1991) Eur. J. Biochem. 197, 563–576.
[2] Chomyn, A., Mariottini, P., Cleeter, M.W.J., Ragan, C.I., Matsuno-Yagi, A., Hatefi, Y., Doolittle, R.F. and Attardi, G. (1985) Nature 314, 592–597.
[3] Chomyn, A., Cleeter, M.W.J., Ragan, C.I., Riley, M., Doolittle, R.F. and Attardi, G. (1986) Science 234, 614–618.
[4] Walker, J.E., Fearnley, I.M., Lutter, R., Todd, R.J. and Runswick, M.J. (1990) Phil. Trans. R. Soc. Lond. B. 326, 367–378.
[5] Weiss, H., Leonard, K. and Neupert, W. (1990) Trends Biochem. Sci. 15, 178–180.
[6] Yagi, T. (1991) J. Bioeng. Biomembr. 23, 211–225.
[7] Berks, B.C. and Ferguson, S.J. (1991) Biochem. Soc. Trans. 19, 581–588.
[8] Zannoni, D. and Ingledew, W.J. (1983) FEMS Microbiol. Lett. 17, 331–334.
[9] Meinhardt, S.W., Kula, T., Yagi, T., Lillich, T. and Ohnishi, T. (1987) J. Biol. Chem. 262, 9147–9153.
[10] Albracht, S.P.J., Van Verseveld, H.W., Hagen, W.R. and Kalkman, M.L. (1980) Biochim. Biophys. Acta 593, 173–186.
[11] Yagi, T. (1986) Arch. Biochem. Biophys. 250, 302–311.
[12] Xu, X., Matsuno-Yagi, A. and Yagi, T. (1991) Biochemistry 30, 8678–8684.
[13] Xu, X., Matsuno-Yagi, A. and Yagi, T. (1991) Biochemistry 30, 6422–6428.
[14] Marmur, J. (1961) J. Mol. Biol. 3, 208–218.
[15] Dupuis, A., Skehel, J.M. and Walker, J.E. (1991) Biochemistry 30, 2954–2960.
[16] Pilkington, S.J., Skehel, J.M. and Walker, J.E. (1991) Biochemistry 30, 1901–1908.
[17] Wu, L.-F. and Saier, J. (1991) Res. Microbiol. 142, 943–949.
[18] Armstrong, G.A., Alberti, M., Leach, F. and Hearst, J.E. (1989) Mol. Gen. Genet. 216, 254–268.
[19] Kohchi, T., Shirai, H., Fukuzawa, H., Sano, T., Komano, T., Umesono, K., Inokuchi, H., Ozeki, H. and Ohyama, K. (1988) J. Mol. Biol. 203, 353–372.
[20] Anderson, S., Bankier, A.T., Barell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.D., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1981) Nature 290, 457–465.
[21] Earley, F.G.P., Patel, S.D., Ragan, C.I. and Attardi, G. (1987) FEBS Lett. 219, 108–113.
[22] Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matusbayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. and Sugiura, M. (1986) EMBO J. 5, 2043–2049.
[23] Fearnley, I.M., Runswick, M.J. and Walker, J.E. (1989) EMBO J. 8, 665–672.
[24] Berger, S., Ellersiek, U. and Steinmüller, K. (1991) FEBS Lett. 286, 129–132.
[25] Takahashi, Y., Shonai, F., Fujita, Y., Kohchi, T., Ohyama, K. and Matsubara, H. (1991) Plant Cell Physiol. 32, 969–981.