Volume 280, number 2, 387–392 FEBS 09536 March 1991

ADONIS 001457939100246M

The δ subunit of the chloroplast ATPase is plastid-encoded in the diatom *Odontella sinensis*

Peter G. Pancic[1], Heinrich Strotmann[1] and Klaus V. Kowallik[2]

[1]*Institut für Biochemie der Pflanzen* and [2]*Institut für Botanik der Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, D 4000 Düsseldorf, Germany*

Received 14 January 1991

A 5.2 kb *Pst*I restriction fragment containing the atpA gene cluster of the plastic genome of the centric diatom *Odontella sinensis* was cloned. Sequencing revealed a reading frame of 561 bp separating the genes atpF and atpA, which is preceded by a putative ribosome binding site. The third nucleotide of the codon for the last amino acid of atpF is the first nucleotide of the initiation codon of the 561 bp reading frame. The amino acid sequence deduced from the nucleotide sequence of this gene (atpD) is colinear with δ subunits of different F_oF_1-ATPases and shows an overall sequence homology of up to 35% when compared with the sequences of cyanobacteria and *Cyanophora paradoxa*. The results are discussed in context with the evolution of chloroplasts of the chlorophyll-*a* + *b* and -*a* + *c* lineages, respectively.

CF_oCF_1-ATPase; δ Subunit; Nucleotide sequence; Amino acid sequence; Chloroplast genome; Diatom; *Odontella sinensis*

1. INTRODUCTION

F_o-F_1-type ATPases of the plasma membrane of eubacteria [1], the inner membrane of mitochondria [2], and the thylakoid membrane of cyanobacteria and chloroplasts [3] couple transmembrane proton translocation with the reversible formation of ATP from ADP and phosphate, Structure, subunit stoichiometry and composition as well as amino acid sequences of the subunits are quite similar in F_o-F_1-ATPases from different sources [4].

The multimeric complex is subdivided into the peripheral F_1 part, containing the catalytic centers for ATP formation [5] and the transmembrane sector F_o which acts as a proton channel. F_1 is composed of 5 subunits (α, β, γ, δ and ε), F_o consists of 3 subunits in eubacteria, 4 subunits in chloroplasts, and at least 6 subunits in mitochondria [4,6].

In *E. coli* the genes for all subunits are tightly linked and constitute a single transcriptional unit [7]. In cyanobacteria they are arranged in two clusters, the atpB gene cluster containing the genes for subunits β (atpB) and ε (atpE) and the atpA gene cluster including all F_o genes together with atpA and atpD, coding for subunits α and δ, respectively. The gene for subunit γ (atpC) may be attached to the atpA gene cluster or isolated, depending on the organism [8–10].

For eukaryotic organellar F_oF_1-ATPases a dual genetic origin has been found. In chloroplasts of higher plants the subunits γ, δ and II are transcribed from nuclear genes and translated on cytoplasmatic ribosomes. Subsequently, the products are imported into the plasmid compartment [11]. The genes for these three subunits are also missing in the chloroplast genomes of the chlorophyll-*a* + *b*-containing algae *Euglena gracilis* [12], *Chlamydomonas reinhardtii* [13] and *Chlamydomonas moewusii* [14]. However, in *Chlamydomonas* the linear array of the chloroplast ATPase genes is scrambled due to extensive intramolecular rearrangements.

Little is known about the organization of ATPase genes in other algal lineages. Among those the Chromophyta deserve particular interest because of the origin of their chloroplasts. In contrast to the Chlorophyta, chromophytic plastids are considered to have evolved from eukaryotic rather than prokaryotic cells [15,16]. As the nucleus of the putative eukaryotic endosymbiont has completely disappeared in most chromophytes, a secondary (eukaryotic/eukaryotic) transfer of genes must have occurred, including those that have previously been transferred from the genome of the evolving chloroplast to the nucleus of the first host cell. The extensively rearranged plastid genomes of chromophytic algae may result from these secondary endocytoses [17].

In this context we were particularly interested in the location of the coding sites for subunits γ, δ and II of chromophytic chloroplast ATPases. Here we report on a plastid gene of the diatom *Odontella sinensis*, which shows sequence similarity to atpD coding for subunit δ. As in eubacteria including cyanobacteria this gene is located in the atpA gene cluster and flanked by the genes atpF and atpA.

Correspondence address: P.G. Pancic, Institut für Biochemie der Pflanzen der Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, D 4000 Düsseldorf, Germany

Published by Elsevier Science Publishers B.V.

2. MATERIALS AND METHODS

2.1. *Isolation of plastid DNA*

Plastid DNA was isolated from purified chloroplasts of *Odontella sinensis* cultivated as described [18]. The cells were harvested by filtration through a 50 μm gauze, washed several times with isotonic NaCl solution (3,5%, w/v) and subsequently disrupted in ice-cold isolation medium (0.4 M mannitol; 50 mM Tris-HCl pH 8.0; 30 mM Na_2EDTA, 4 mM $MgCl_2$; 0.1% BSA; 1 mM mercaptoethanol) using a glass potter homogenizer. The homogenate was filtered through a 20 μm gauze and the filtrate centrifuged for 5 min at 1500 × *g* in order to remove cell debris. The supernatant which essentially contained morphologically intact plastids was then centrifuged for 10 min at 3500 × *g*. The pelleted plastids were washed in modified isolation medium lacking BSA and DTT and centrifuged again. The plastids were transferred into lysis buffer (50 mM Tris-HCl, pH 8.3; 100 mM Na_2 EDTA; 50 mM NaCl; 0.5% SDS; 0.7% lauroyl sarcosinate; 1 mg/ml Proteinase K), and the lysate was phenolized according to standard protocols. The DNA was precipitated by ethanol, dried, redissolved in Tris-EDTA buffer (10 mM/1 mM, pH 7.5) and purified in a CsCl/ethidium bromide density gradient using a vertical rotor at 70 000 rpm for 5 h. The DNA was stored in Tris-EDTA buffer at a final concentration of 0.5 μg/μl.

2.2. *Cloning and sequencing of the Odontella atpD-gene*

Southern hybridization experiments revealed that an atpA gene probe from spinach chloroplasts cross-reacted with a 5.2 kb *PstI* restriction fragment (P5). This fragment was electro-eluted from agarose gels and cloned into pUC-18 vector using *E. coli* strain DH5α as host. The clones were screened by *PstI* digestion, and the insert checked by Southern hybridization using the spinach atpA gene as a probe. The resulting clone (pOsP5) was digested with *pstI–XbaI* and yielded two subfragments of 2.4 and 2.8 kb, respectively. The 2.4 kb fragment (pOsP5X2) contained most of the atpA gene. Nested deletions of this fragment were performed using the exonuclease III/mung bean nuclease enzymes from Boehringer. Klenow fill-in reactions prior to religation enhanced the number of deleted clones. The clones were sequenced according to the dideoxy chain termination method [19] using the Pharmacia T7 sequencing kit.

For DNA sequence analysis and translation we used the computer program 'DM' [20]. Multiple alignments were carried out using the program 'Clustal' [21]. The program 'SOAP' [22] served to calculate hydrophobicity plots.

3. RESULTS

Sequence analysis of the fragment pOsP5X2 revealed 1452 nucleotides with a striking similarity to atpA genes

```
ATTATCCTATTAGTTTTAAAACGTACAGTAGCTCGCGCTCAGCAAACTTTTGGTCCAAAAGAAAG
 I  I  L  L  V  L  K  R  T  V  A  R  A  Q  Q  T  F  G  P  K  E  R

AGCAACAGCATTAATTACTGAAACAATTAATAAATTAGAAGGAGATTTGTTATGAGTATAAATCC
  A  T  A  L  I  T  E  T  I  N  K  L  E  G  D  L  L Ter atpF
                                          atpD  --> M  S  I  N  P

                                           50
TTTAGCTTCAAAAATTGCAGCTCCTTATGCACGTGCTTTGTTTGACTTCTCAGTTGATCAAAATC
  L  A  S  K  I  A  A  P  Y  A  R  A  L  F  D  F  S  V  D  Q  N

                     100
TTATGCATCAAATTACTGCTGATTTTCAGAATTTAGAAGTCTTTTTAAATAAAACACCTGATTTA
L  M  H  Q  I  T  A  D  F  Q  N  L  E  V  F  L  N  K  T  P  D  L

      150                                                200
ACAGAATATTTAAGTAATCCTCTTATTAGTGCAAAATCGAAAGAGGAAGTTTTAAATAAAACTTT
 T  E  Y  L  S  N  P  L  I  S  A  K  S  K  E  E  V  L  N  K  T  L

                                                250
AAAATCGCAAATAAACAAAGAAACTTTTAAATTTTTAATTGTTTTAGTAAATCGAAGTAGAATTA
  K  S  Q  I  N  K  E  T  F  K  F  L  I  V  L  V  N  R  S  R  I

                              300
ATTTGTTAGAACCCATTATTGCTAGTTATCTAAATCTAGTTTATAATGCAGCTTCTGTTAAAATG
N  L  L  E  P  I  I  A  S  Y  L  N  L  V  Y  N  A  A  S  V  K  M

             350                                             400
ATTGAAGTTTCTACTGCTTATGCATTTACGAATTTGCAAAAAAATACTTTAATTAAAAAATTAAA
 I  E  V  S  T  A  Y  A  F  T  N  L  Q  K  N  T  L  I  K  K  L  K

                                                 450
AGAATTAACAAACGCAAGAGAAATTCGCTTAGTAATTACTGTTGATTCAAGTCTTATTGGCGGTT
  E  L  T  N  A  R  E  I  R  L  V  I  T  V  D  S  S  L  I  G  G

                                    500
TTTTAATTAAAACTAATTCAAAAGTACTTGATTTTACCATTAAAAATCAATTACAAAAATTAGCT
F  L  I  K  T  N  S  K  V  L  D  F  T  I  K  N  Q  L  Q  K  L  A

                  550
AAACATTTAGATAGTGTTTTAGAAATTTAAAAATAACTAAATTCTTTTATTAACTTTATTATTTT
 K  H  L  D  S  V  L  E  I Ter atpD

AAAAATAATACAATGATAAATATTCGTCCAGACGAAATTAGTAGTATTATCCGTGAACAAATTGA
    atpA  --> M  I  N  I  R  P  D  E  I  S  S  I  I  R  E  Q  I  E

ACAATATGATCAAGATGTTAAAGTAGATAATATTGGTACTGTATTACAAGTAGGTGATGGTATTG
  Q  Y  D  Q  D  V  K  V  D  N  I  G  T  V  L  Q  V  G  D  G  I
```

Fig. 1. Nucleotide sequence and deduced amino acid sequence of the gene atpD of the plastid genome of *Odontella sinensis* together with the upstream (atpF) and downstream (atpA) sequences. Numbers above the atpD sequence start from the presumptive initiation codon. A putative ribosome binding site is underlined, start and termination codons are in italics.

Fig. 2. Aligned amino acid sequences deduced from atpD genes of *Odontella sinensis*, *Cyanophora paradoxa* (D.A. Bryant, V.L. Stirewalt and M.B. Annarella, unpublished results), *Anabaena* PCC 7120 [9], *Synechocystis* 6301 [8], *Rhodopseudomonas blastica* [28], *Spinacia oleracea* [11], *E. coli* [29] and of bovine mitochondria (OSCP) [2]. Amino acid residues identical to the *Odontella* protein are boxed. Asterisks mark amino acids that occur at the same position in all organisms.

	1	2	3	4	5	6	7	8	9	10	11	12
1. *Odo.sin*	100											
2. *Cya.par.*	35	100										
3. *Ana.7120*	32	37	100									
4. *Syn.6301*	31	36	46	100								
5. *Rsp.rub.*	23	22	28	25	100							
6. *Rps.bl.*	24	24	26	24	41	100						
7. *Spin.ol.*	25	27	26	25	24	23	100					
8. *PS3*	18	18	25	28	23	23	22	100				
9. *B.mega*	17	19	26	22	21	21	20	44	100			
10. *E.coli*	21	23	25	20	19	24	21	23	18	100		
11. *Bov.OSCP*	21	22	18	19	25	23	21	23	17	23	100	
12. *Ipo.bat.*	12	17	15	15	19	16	17	17	16	17	30	100

Fig. 3. Sequence similarity matrix of δ subunits from different sources. In addition to the sequences listed in Fig. 2 the δ subunits of *Rhodospirillum rubrum* [30], PS3 [31], *Bacillus megaterium* [23] and of mitochondria of sweet potatoe [32] are included. Similarity is given in percentage identical amino acids.

from different sources (data not shown). About 70–80% of the deduced amino acid residues are identical with those from α subunits of land plant chloroplasts and cyanobacteria. 5′ to the atpA gene, and separated by 47 bp, is a reading frame of 561 bp with a coding capacity for 187 amino acids. The start codon 5′ATG3′ of this reading frame overlaps with the codon for the last amino acid and the termination codon 5′TGA3′ of a gene that was identified as atpF coding for CF_0 subunit I. A G+A-rich sequence presumably containing a ribosome binding site was found around 10 bp upstream of the start codon of the 561 bp reading frame. Fig. 1 shows the nucleotide sequence of this reading frame together with the adjacent 5′atpA and 3′atpF sequences and the deduced amino acid sequence. The molecular mass of the 561 bp gene product was calculated as 21.1 kDa.

Comparison of the amino acid sequence with known sequences of ATPase polypeptides revealed that the gene product resembles subunit δ from different sources (Fig. 2). The sequence of the *Odontella* gene product is co-linear with other prokaryotic sequences and does not contain the three amino acid insertion near the C-terminus of the pinach protein.

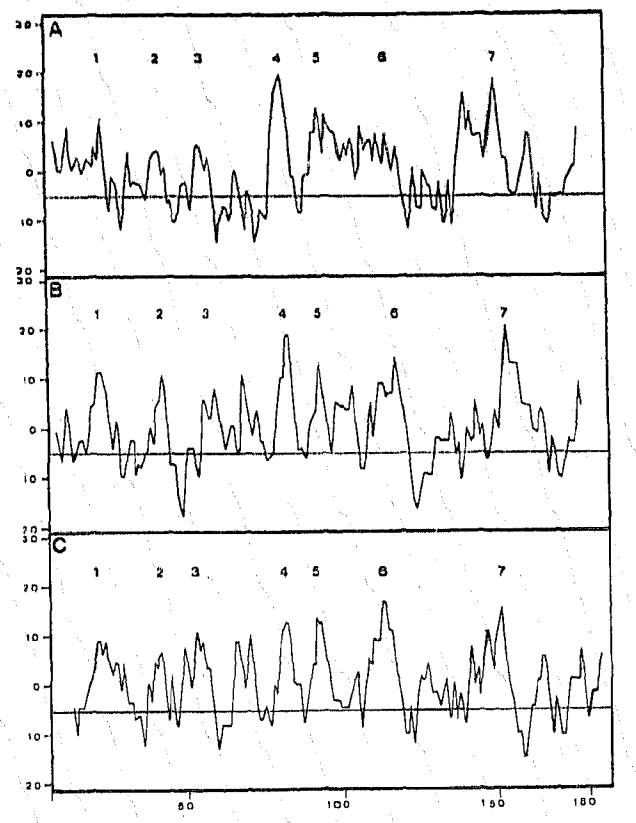

Fig. 4. Hydropathy plots of the δ subunits of (A) *Odontella sinensis*, (B) *Spinacia oleracea* [11] and (C) *Anabaena* PCC 7120 [9]. Calculations were performed according to [22] with a window of 7 amino acids. Characteristic and comparable hydrophobic regions are numbered.

Although the overall sequence homology among these proteins is relatively poor (Fig. 3). The N- and C-terminal regions exhibit some degree of sequence conservation. Seven out of nine amino acid residues that are identical in other δ subunits are also present in the deduced sequence of *Odontella*. Sequence conservation is increased when only the polypeptides of organisms performing oxygenic photosynthesis are compared. The overall sequence homology is highest between the gene product of *Odontella* and the δ subunits of cyanobacteria and *Cyanophora paradoxa* (Fig. 3).

The hydropathy plots of the *Odontella* gene product and of δ subunits from spinach chloroplasts [11] and the blue-green alga *Anabaena* PCC 7120 [9] are compared in Fig. 4. They indicate that the *Odontella* protein — as the δ subunits of other ATPases — most likely is not a membrane-anchored protein. This view is confirmed by calculations of secondary structures using published computer programs (data not shown). Nevertheless, several hydrophobic regions can be discerned which are located in similar positions of the three protein sequences.

4. DISCUSSION

In chlorophyll-*a* + *b*-containing eukaryotes examined so far, the genes for subunits γ, δ and II of the chloroplast ATPase were shown to reside in the nuclear genome. Accordingly, these genes are missing in the corresponding region of the plastid chromosome when compared with the eubacterial operons [10]. In the chromophyte *Odontella sinensis*, however, at least the investigated part of the atpA gene cluster including the gene atpD resembles the eubacterial gene order (Fig. 5).

The identity of the 561 bp reading frame in the plastid genome of *Odontella* as the gene coding for subunit δ is based on (1) its coding capacity which is equivalent to other prokaryotic and eukaryotic atpD genes, (2) its sequence similarity with δ subunits of ATPases from other sources, especially from cyanobacteria, and (3) its hydrophobicity pattern showing similarities to those of other δ subunits. In addition, the atpD reading frame of *Odontella* overlaps in the same way with atpF as in *Cyanophora paradoxa* (D.A. Bryant, V.L. Stirewalt and M.B. Annarella, unpublished results), *Anabaena* PCC 7120 [9] and *Bacillus megaterium* [23] (Fig. 6). Such overlapping genes are extremely rare in cyanobacteria and chloroplasts. A well-known example are the genes atpB and atpE which overlap in certain land plants by 1 bp, but are separated in others [3]. This atpBE overlap which apparently resembles a late evolutionary event in land plant lineages, does neither exist in cyanobacteria [8,24] nor in *Odontella* (Pancic, et al., unpublished results) or *Dictyota* (Leitsch and Kowallik, in preparation). A similar prokaryotic feature is found in the psbDC operon of another chromophyte, *Vaucheria bursata*, exhibiting a 14 bp overlap (Kowallik et al.,

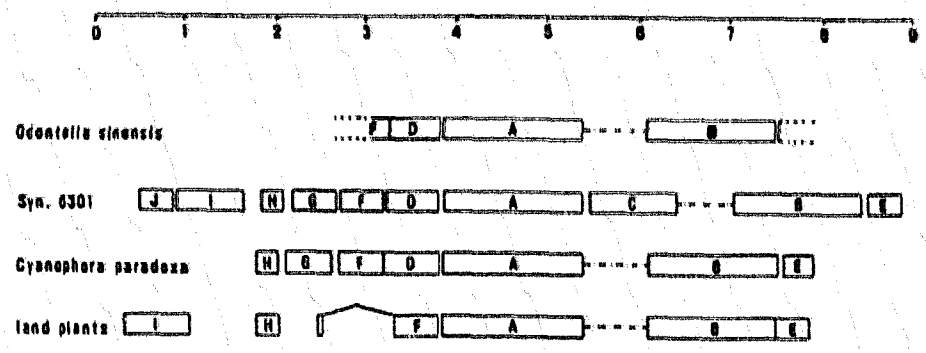

Fig. 5. Arrangement of the genes for F_0F_1-ATPases in *Anabaena*, in cyanelles of *Cyanophora* and in chloroplasts of *Odontella sinensis* and higher plants. A bent line connects the two exons of the land plant atpF gene. Overlapping genes are separated by a single line, a dashed line separates the two gene clusters. The scale is in kb.

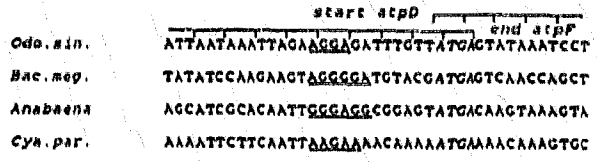

Fig. 6. Nucleic acid sequences showing the 4 bp overlap (in italics) between atpF and atpD of *Odontella sinensis*, *Bacillus megaterium* [23], *Anabaena* PCC 7120 [9] and *Cyanophora paradoxa* (D.A. Bryant, V.L. Stirewalt and M.B. Annarella, unpublished results). Possible ribosome binding sites are underlined.

unpublished results) as in *Synechococcus* PCC 7942 [25], where these genes share 50 bp in land plant chloroplasts [26]. These results suggest a close relationship between cyanobacteria and chromophytic chloroplasts.

The unexpected gene composition manifested in the atpA gene cluster of *Odontella* points to the question how chromophytic and chlorophytic plastids have evolved. There is now do doubt that both types of plastids originated from cyanobacterial ancestors although chromophytes may have experienced different kinds of secondary (eukaryotic/eukaryotic) endocytoses [27]. The existence of a gene for subunit δ in the plastid chromosome of *Odontella* at the same position as in cyanobacteria and the cyanelles of *Cyanophora paradoxa* suggests that the transfer of this gene into the nuclear genome has occurred within the chlorophyll-$a+b$ lineage only. This suggestion is supported by the finding that the same gene arrangement of the *Odontella* atpA gene cluster was found in the brown alga *Dictyota dichotoma* (Kuhsel et al., in preparation).

The identification of a plastid gene in the Chromophyta that is nuclear in chlorophyll-$a+b$-containing organisms now renders the search for the coding sites of atpC and atpG in chromophytes attractive. We may then perhaps answer the question as to whether the transfer of the genes atpC, atpD, and atpG results from a single event which may be unique for the chlorophyll-$a+b$ lineage or whether it reflects a multistep evolutionary process.

Acknowlegements: The authors thank Prof. D.A. Bryant, Pennsylvania State University, for the kind permission to show unpublished data and Prof. R.G. Herrmann, University of Munich, for providing the spinach atpA gene probe. Furthermore, we thank Thorsten Friedrich, University of Düsseldorf, for help in secondary structure calculations. This work was supported by the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 189) and the Fonds der Chemischen Industrie.

REFERENCES

[1] Senior, A.E. (1988) Physiol. Rev. 68, 177–132.
[2] Walker, J.E., Fearnley, I.M., Gay, N.J., Gibson, B.W., Northrop, F.D., Powell, S.J., Runswick, M.J., Saraste, M. and Tybulewicz, V.L.J. (1985) J. Mol. Biol. 184, 677–701.
[3] Hudson, G.S. and Mason, J.G. (1988) Photosynth. Res. 18, 205–222.
[4] Walker, J.E., Cozens, A.L., Dyer, M.R., Fearnley, I.M., Powell, S.J. and Runswick, M.J. (1987) Biochem. Soc. Trans. 15, 104–106.
[5] Xue, Z., Zhou, J.M., Melese, T., Cross, R.L. and Boyer, P.D. (1987) Biochemistry 26, 3749–3753.
[6] Strotmann, H. and Bickel-Sandkötter, S. (1984) Annu. Rev. Plant Physiol. 35, 97–120.
[7] Futai, M. and Kanazawa, H. (1983) Microbiol. Rev. 47, 285–312.
[8] Cozens, A.L. and Walker, J.E. (1987) J. Mol. Biol. 194, 359–383.
[9] McCarn, D.F., Whitaker, R.A., Alam, J., Vrba, J.M. and Curtis, S.E. (1988) J. Bact. 170, 3448–3458.
[10] Palmer, J.D. in: Cell culture and Somatic Cell Genetics in Plants, Vol. 7 (Vasil, I.K., ed) in press.
[11] Hermans, J., Rother, C., Steppuhn, J. and Herrmann, R.G. (1988) Plant. Mol. Biol. 10, 323–330.
[12] Hallick, R.B. and Buetow, D.E. (1989) in: The Biology of *Euglena* (Buetow, D.E., ed) Vol. IV, pp. 351–414, Academic Press, New York.
[13] Woessner, J.P., Gilham, N.W. and Boynton, J.E. (1987) Plant Mol. Biol. 8, 151–158.
[14] Turmel, M., Lemieux, B. and Lemieux, C. (1988) Mol. Gen. Genet. 214, 412–419.
[15] Gibbs, S.P. (1981) Ann. NY Acad. Sci. 361, 193–208.
[16] Whatley, J.M. and Whatley, F.R. (1981) New Phytol. 87, 233–247.

[17] Kowallik, K.V. (1989) In: The Chromophyte Algae. Problems and Perspectives, The Systematics Association, Vol. 38 (Green, J.C., Leadbeater, B.S.C. and Diver, D.L., eds) pp. 101–124, Clarendon, Oxford.
[18] Pancic, P.G., Kowallik, K.V. and Strormann, H. (1990) Bot. Acta 103, 274–280.
[19] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[20] Wilbur, W.J. and Lippman, D.J. (1983) Proc. Natl. Acad. Sci. USA 80, 726–730.
[21] Higgins, D.G. and Sharp, P.M. (1988) Gene 73, 237–244.
[22] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[23] Brusilow, W.S.A., Scarpetta, M.A., Hawthorne, C.A. and Clark, W.P. (1989) J. Biol. Chem. 264/3, 1528–1533.
[24] Curtis, S.E. (1987) J. Bact. 169, 80–86.
[25] Golden, S.S. and Stearns, G.W. (1988) Gene 67, 85–96.
[26] Ali, J., Morris, J., Westhoff, P. and Herrmann, R.G. (1984) Curr. Genet. 8, 597–606.
[27] Whatley, J.M. and Whatley, F.R. (1984) In: Encyclopedia of Plant Physiology (Linsgens, H.F. and Heslop-Harrison, J., eds) pp. 18–58, Springer, Berlin.
[28] Tybulewicz, V.L.J., Falk, G. and Walker, J.E. (1984) J. Mol. Biol. 179, 185–214.
[29] Walker, J.E., Gay, N.J., Saraste, M. and Eberle, N. (1984) Biochem. J. 224, 799–815.
[30] Falk, G., Hampe, A. and Walker, J.E. (1985) Biochem. J. 228, 391–407.
[31] Ohta, S., Yohda, M., Ishizuka, M., Hiruta, H., Hamamoto, T., Otawara-Hamamoto, Y., Matsuda, K. and Gawa, Y. (1988) Biochim. Biophys. Acta 933, 141–155.
[32] Kimura, T., Takeda, S., Asahi, T. and Nakamura, K. (1990) J. Biol. Chem. 265, 6079–6085.