ELSEVIER

Biochimica et Biophysica Acta 1237 (1995) 95–98

BBA Biochimica et Biophysica Acta

Short Sequence-Paper

Isolation of the *vma-4* gene encoding the 26 kDa subunit of the *Neurospora crassa* vacuolar ATPase

Emma Jean Bowman, Alicia Steinhardt, Barry J. Bowman *

Department of Biology, Sinsheimer Laboratories, University of California, Santa Cruz, CA 95064, USA

Received 21 December 1994; revised 10 April 1995; accepted 13 April 1995

Abstract

We have isolated the *vma-4* gene, which encodes a 25 746 Dalton subunit of the vacuolar ATPase, from *Neurospora crassa*. The gene contains two introns and was mapped to the left arm of linkage group I. Comparison of the predicted amino acid sequence with homologous proteins from *Saccharomyces cerevisiae, Manduca sexta*, and *Bos taurus* showed only 25% sequence identity. However, computer-assisted predictions of secondary structures gave similar results for all four proteins. Analysis of the sequence and the available biochemical data indicated that the *vma-4* gene product may play the same structural role in the vacuolar ATPase as does the γ-subunit in F-type ATPases.

Keywords: Vacuolar ATPase; ATPase, F-type; Vacuole; Proton pump; Gamma subunit; (*Neurospora crassa*)

The vacuolar ATPase is a large, complex enzyme which functions as a proton pump in many different types of cellular membranes [1]. In *Neurospora crassa* this enzyme may constitute 10–20% of the protein in the vacuolar membrane [2]. By pumping protons into the interior of the vacuole the vacuolar ATPase acidifies the organelle and also generates an electrochemical gradient used to drive the transport of small molecules [3].

The polypeptides making up the vacuolar ATPase are organized into two sectors. The integral membrane sector, called Vo, is composed of at least four types of polypeptides, 100 kDa, 40 kDa, 20 kDa and 16 kDa in size, in a stoichiometry of 1:1:1:6 [4]. Protruding from the membrane is a peripheral sector called V1. The *N. crassa* V1 is composed of at least five different polypeptides, 67 kDa, 57 kDa, 48 kDa, 30 kDa and 17 kDa [5], in a probable stoichiometry of 3:3:1:1:1 by comparison to the bovine coated vesicle enzyme [4]. Comparative analyses of the overall structures of the enzymes and of the primary sequences of several subunits strongly suggest that the vacuolar ATPase is homologous to the F-type ATPase found in mitochondrial, chloroplast and eubacterial membranes [6]. For only three subunits, however, is there a clear correspondence between F- and V-type ATPases. The 67 kDa subunit (encoded by *vma-1*) is the homolog of the β-subunit of F-type ATPase, the 57 kDa subunit (*vma-2*) is the homolog of the α-subunit [6], and the 16 kDa Vo subunit (*vma-3*) is the homolog of the F-type 'c' subunit [7]. The degree of sequence conservation among other subunits is low, even among F-type ATPases, and the limited data available for other vacuolar ATPase subunits have not allowed identification of F-type homologs.

In this paper we report the sequence of the gene that encodes the 30 kDa subunit of the vacuolar ATPase of *N. crassa*. All vacuolar ATPases appear to have a subunit of approx. 30 kDa, about the same size as the γ-subunit in F-type ATPases [1,8–11]. We wanted to identify conserved regions within the 30 kDa subunit and to see if the primary structure and the predicted secondary structure of the 30 kDa subunit showed similarity to the γ-subunit of the F-type ATPase.

Isolation of the vma-4 gene. The V1 sector of the vacuolar ATPase was prepared as described [5]. The 30 kDa subunit was isolated from polyacrylamide gels and digested with trypsin [12]. After separation by HPLC [12], six tryptic peptides were sequenced and used to design oligonucleotides for use in the polymerase chain reaction. The 5′ oligonucleotide AGAT(C/T)CA(G/A)AT(C/T)-AA(G/A)GC(C/T)GA(C/T)GA(G/A)GA and the antisense of the 3′ oligonucleotide GA(C/T)GA(G/A)AT-(C/T)TT(C/T)GA(G/A)GC(C/T)GC(C/T)TCCGC gave a 215 bp PCR product, which was used to screen an

* Corresponding author. E-mail: bowman@orchid.ucsc.edu. Fax: +1 (408) 4593139.

SSDI 0005-2736(95)00108-5

N. crassa cDNA library in the λZAP vector (obtained from Sachs, M., Oregon Graduate Institute). We obtained a partial cDNA fragment (truncated at an *Eco*RI restriction site near the 5′ end) that encoded five of the tryptic peptides we had sequenced (Fig. 1) and showed sequence similarity to the *VMA4* gene from yeast [13]. The cDNA was then used to screen a genomic library [14]. Four positive cosmid clones (14-5-H, 15-5-C, 21-7-E, and 29-3-H) were selected. From these clones a 2.9 kb *Pst*I restriction fragment was subcloned, sequenced, and found to encode the complete gene, named *vma-4*. By using a 484 bp *Eco*RV/*Eco*RI fragment of the genomic DNA to

```
   1 AGAGCGACTCGTGAGTTCAGTTGGCTGAGTGATGTCCACCACCACCACCACCACCAGCACCACCACCTAGCTGGAGGTTCAGTTCGTTGC
  91 CCAGGGCCGGGGTTGCTGTTGCCAGCCTTTCTGGCTCCATCCCAGTGCCAAGCCAGCATCATCACCCTTCCCCAAGCTGCGTCAACGGGC
 181 AACAACAACATCCATTAGTCGACGTCGTTTGACGATATCCTCCGTCTAAACTAAATACTACCAGCAACCGCGATACCCTTTGAGCCCTGT
                                 start of cDNA
 271 TTTCCACTTTTGATAGATCGCCATCCCCTCAACGTCCCTACACCTCACGTCTTCGTCGTCTTACACTCGACATCGATCCAAGTTAGCTCT

 361 GCCCGCCACGAGCTTCCCCTCCCAAGGACGACGACGACATCACATCCTCCATCACCGTCCTTCGGAGACCCGAACCACCACCACCACCAC

 451 CACCACAACACCTATTCCCTAGAGTGCCTTCATTAACACCCGCCAAAATGTCACAAGTTCACGCCTTGTCCGACGATCAGgtagagccaa
                                                    M  S  Q  V  H  A  L  S  D  D  Q

 541 ttccgtccagtccagataaccaaaccatgcaatgccagctaaccaaagaaccacgctgcccaccaaccagGTCGGACAAGAGCTCCGCAA
                  Intron #1                                                    V  G  Q  E  L  R  K
                                                        5' PCR Primer      EcoRI
 631 GATGACGGCCTTCATCAAGCAAGAAGCCGAGGAAAAGGCGCGCGAGATCCAGATCAAGGCCGACGAGGAATTCGCCATTGAAAAGTCCAA
      M  T  A  F  I  K  Q  E  A  E  E  K  A  R  E  I  Q  I  K  A  D  E  E  F  A  I  E  K  S  K

 721 GCTGGTCCGCCAGGAGACGGACGCCATCGACTCGGCCTACGCCAAAAAGTTCAAGCAGGCCCAGATGTCGCAGCAGATCACGCGGTCGAC
      L  V  R  Q  E  T  D  A  I  D  S  A  Y  A  K  K  F  K  Q  A  Q  M  S  Q  Q  I  T  R  S  T
                                                                3' PCR Primer
 811 AATGGCCAACAAGACGCGACTCCGGGTCCTGGGCGCGCGCCAGGAGCTGCTGGACGAGATCTTTGAGGCCGCCAGCGCCCAGCTCGGCCA
      M  A  N  K  T  R  L  R  V  L  G  A  R  Q  E  L  L  D  E  I  F  E  A  A  S  A  Q  L  G  Q

 901 AGCCACCCACGACCTGGGCCGCTACAAGGACATCCTCAGGGACTTGATCCTCGAGGGCTTCTACGCCATGAACGAGCCGGAGCTGGTGAT
      A  T  H  D  L  G  R  Y  K  D  I  L  R  D  L  I  L  E  G  F  Y  A  M  N  E  P  E  L  V  I

 991 CCGCGCGCGCCAGGCGGATTATGACGCGGTCAGGGAGGCGGCGGGGTGGGCGAGCGCCCAGTACAAGCACAAGACGGATAAGGATGTCAA
      R  A  R  Q  A  D  Y  D  A  V  R  E  A  A  G  W  A  S  A  Q  Y  K  H  K  T  D  K  D  V  K

1081 GGCGACGATTGATGCGGAGAATCCCGTTCCTGAGGGGAGgtaggttatacatacacacaccatgcttcgggctgtggtgaagaggaagaa
      A  T  I  D  A  E  N  P  V  P  E  G  S                                  Intron #2

1171 gaagaagagatgatgctaacaagagttgatatacagCGCCGGTGGCATCATCATCGTTGGCGGAAACGGCAAGATCGATATCGACAACAC
                                                A  G  G  I  I  I  V  G  G  N  G  K  I  D  I  D  N  T

1261 CTTCGAGGCCAGGTTGACCCTGCTTAAGGACAGCGCTCTCCCGGCGATGCGCAAGGCTCTGTTCGGTGAAAACCCCAACCGCAAGTTCTT
      F  E  A  R  L  T  L  L  K  D  S  A  L  P  A  M  R  K  A  L  F  G  E  N  P  N  R  K  F  F

1351 CGATTAAATGCACCACTTTTTTGGAAGGCTGCGAATGCGCAGAAGGCCTACAACAAAGCTACGGCACCTGGTGTCCAGGAAAGTGAGAAA
      D  end

1441 GGGGTCCTGTGGACAAATCTCCAGAGTGAGACGAAAGCCGGAGGTGTACACCAGAGCGTTGGGTGCTTACTTACTACTCGTTGTTTGATG
1531 ATGTTCAATGACAAAGACGACAAGTTAAGCCAAGAAAAGGTTGGCGAAGGAGTATGGTAGGTGGTTATGGATGATGGACGAATGCATGCA
                                    end of cDNA
1621 TGGAGGATTATCATGTTAGAGTTTTGTATGATACAGTTGTTTTCGCTGTTTTCCCCTCTTGTTTATCGCAAAGCAACTGTTGTTTCACGC
1711 TGTTGGCATCCTTTGAGTTCCTCTAGACTTTGCCAATCTTTGAGAACTGTACGTTTCCTGACATGGCCAACCATTTACCTCGTATCTTAA
1801 CGCTTCTATATTGCCTTTCATTCCATATACCGCAGACGAACAAACAGGCAAAGCAAGCCGAATATACACCGTACCCAATAACCCACAAAG
1891 TTCATCAGTAGTAAACAGAGTTCCTCATAACGCTCAAACACCAAACATGTCCATATATCCCAGTAAAAGCCCCCATTCAAGCCTCCTCCG
1981 TCTTCGCTGCCCCCGCCGTCGTCTTCCCCAAAAACCTCCTCATAAACATCTTTTGCTCCACCCCCTCCCTCTCCCTTCTCTGTTCCCGCT
2071 CCCTCTCCTCCTCACTATCCACGTCCTCCCCCTGCTTCCCATACCCAGTACTACCACCACCCAAGATCCCCGCCCCTCCTCCCGCAGACC
2161 GAATCCCAAAGATATTCGTCCTGCTCACCGGCTTCGTCCTAACCGCCCACTCCAATTTTCTCTCCAAGTCTTTTAGCTTGACTTTGGCTT
2251 CCAATCGAGAAATCTCTGCCGTTGACCCGGATTGGTCGTGTATGTTGGACCAAAGAGTGCGGATGCGGCGGGTGAGGAGGTCGATTTGGC
2341 GGTTATGCGATTTGGTTTCGGGGTTTAGTGTTGTTGTGGTGGATGTTGTTGTTGTTGTTGTTGATTTAGGGAACGGAAGGGAACCAGAAC
2431 CAGAAGAAGAAGAAGAGGGAGGGGGAGGATCGCCGTCGCCGATGGATATGCGTAATAGGTCGGTTTGGGCCGAGAGAACAAAGTAATCGG
2521 CGGACGAGGAAGAACGGGTCGTCTGAAAGTGGGGAGAGATGTCGAGGTAGGAGAGAAGTAAAGGGAAGGAAAAAGTCTGGGAAGGGTCGT
2611 GGAGGGAAAATTCTTGGATGGGGGAAGAGGGGATCTTCTTGGATGATGATGATGATGATGATGATAGAGAAGAAGGGGGGA
2701 AAGAGGGTAAAAGAGAGGTGGAAGCGGAGGCGAAGAGGAAGGCGAGCGCTAGGCGGCGGCGGAGGTTGACTAGTTTCGGGGTTTGGTAGG
2791 GTAGGGCACGGAGGGCGTCCGAGAGGAGGGAGGAGTCCGTGGTTAGGGCGTAGAGACTGCTGCAG
```

Fig. 1. Nucleotide and amino acid sequence of the genomic region encoding *vma-4*. A 2.9 kb *Pst*I restriction fragment was sequenced by the method of Sanger [29]. The region from nt 2270 to the 3′ end was sequenced only in the 3′ to 5′ direction. Both strands were sequenced for nt 1–2270. The 5′ *Pst*I site, approximately 15 bp upstream of the 5′ end, is not shown because of an unresolvable sequence ambiguity in that region. Boxed nucleotides show the region corresponding to the PCR primers used to clone the cDNA. Boxed amino acids are identical to peptide fragments sequenced from the vacuolar ATPase polypeptide.

rescreen the cDNA library, we obtained a second *Eco*RI cDNA fragment derived from the 5′ end of the gene. Fig. 1 shows the sequence of the genomic DNA, the regions corresponding to the cDNA, the derived amino acid sequence, and the amino acid sequences corresponding to the tryptic peptides.

Mapping of the vma-4 gene. The chromosomal location of the gene was determined by analysis of restriction fragment length polymorphisms [15], using a polymorphism for an *Eco*RI site. *Vma-4* mapped to the left arm of linkage group I, near the *Fsr-12* gene. None of the five other identified genes encoding vacuolar ATPase subunits in *N. crassa* are found on this linkage group [16] (and unpublished results).

Characteristics of the vma-4 gene and its transcript. By comparing cDNA and genomic sequences we identified two introns within the coding region. As in many other genes from filamentous fungi, including the *N. crassa* ATPase genes [16,17], the introns were positioned near the ends of the coding region, one after the twelfth codon and the other near the 3′ end (Fig. 1). Also common to genes of filamentous fungi [16,17], we found no apparent TATA element within 300 bps of the 5′ end of the cDNA and no apparent AATAAA signal for polyadenylation. We sequenced the 1200 bp region beyond the polyadenylation site and used it to search GenBank version 85. No sequence with significant similarity was identified. Using the cDNA as a probe of an RNA blot, we found a single size class of transcript, approximately 1.4 kb, similar to the size of the cDNA (1159 bp, excluding the poly(A) tail) (data not shown).

Characteristics of the vma-4 protein. The protein encoded by *vma-4* contains 230 amino acids, has a molecular weight of 25 746 Daltons, and has a p*I* of 5.45. Homologous genes or cDNAs encoding vacuolar ATPase subunits have been reported from three different organisms. Bovine [18] and insect (*Manduca sexta*) [19] subunits are similar to each other with 64% amino acid identity. The yeast subunit [13,20] is significantly less similar, 33–34% identical to the other two. Interestingly, the *vma-4* product from *N. crassa* is equally similar to each of the other three, 40–42%. Overall, the protein is not highly conserved with 24.7% identity among all four organisms.

The *N. crassa* gene is the only one for which intron positions have been reported. The position of the second intron corresponds to a site at which the yeast protein appears to have two extra amino acids when compared to the others (Fig. 2). Conceivably, this could mark the boundary of a domain within the protein.

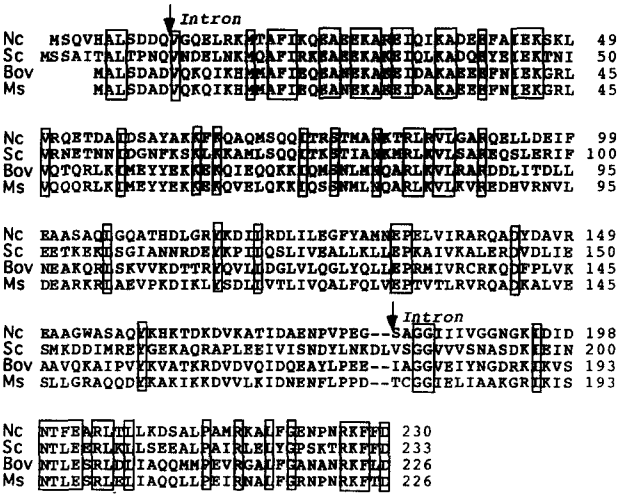

Fig. 2. Alignment of the amino acid sequence of *vma-4* with three homologous proteins. The sequences are from *N. crassa* (Nc, this report), *S. cerevisiae* (Sc, [13]), bovine kidney (Bov, [18]), and *Manduca sexta* (Ms, [19]). Residues identical in all four sequences are boxed. Arrows indicate the position of introns in the protein coding region of the *N. crassa vma-4* gene.

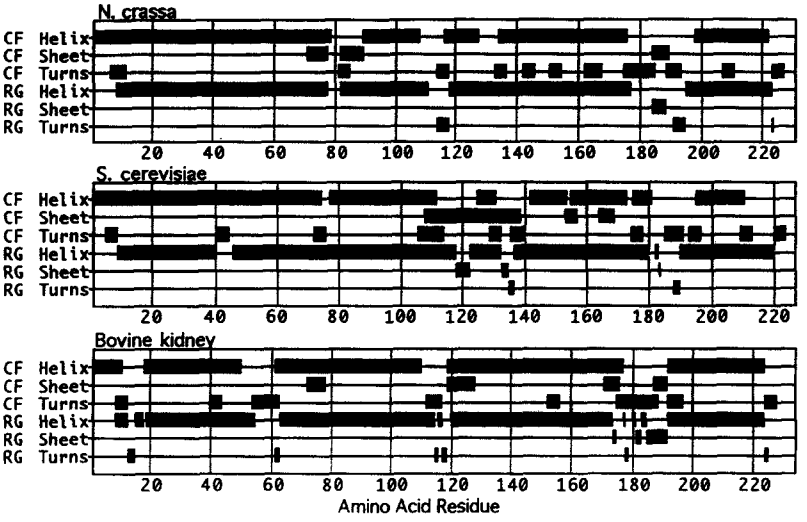

Fig. 3. Predicted secondary structures of the *vma-4* gene product and two homologs. The secondary structures were predicted by the algorithm of Chou and Fasman (CF) [30] or of Robson and Garnier (RG) [21], using the MacVector program (IBI, New Haven, CT, USA).

Although the amino acid sequences of these proteins are not highly conserved, computer-assisted analyses of the 30 kDa proteins gave similar predictions of secondary structures (Fig. 3). Especially when analyzed by the algorithm of Robson-Garnier [21] the proteins are predicted to be predominantly α-helical with conserved turn regions near residues No. 120 and No. 180.

Possible role of the 30 kDa subunit in the vacuolar ATPase. The primary sequence of the 30 kDa subunit is not obviously similar to any subunit of the F-type ATPase. However, several lines of evidence lead us to suggest that the *vma-4* protein may have the same structural role in the vacuolar ATPase as the F1 γ-subunit in the F-type ATPase. In the F-type ATPase the minimal structure that retains significant ATPase activity is $\alpha_3\beta_3\gamma$ [22]. For the bovine coated vesicle vacuolar ATPase, significant catalytic activity has been reported for a complex that contains only four subunits, 70, 58, 40 and 33 kDa [23,24]. Because the 70 and 58 kDa subunits are clearly the homologs of, respectively, the α- and β-subunit in F-type ATPase, one of the other two subunits could be the homolog of the γ-subunit in F-type ATPase. Puopolo et al. [25] reported that the bovine 40 kDa subunit was not essential for activity when integral membrane and peripheral sectors of the ATPase were reassembled from dissociated subunits. In addition, the peripheral sector of the vacuolar ATPase from *N. crassa* does not appear to have a counterpart to the 40 kDa subunit observed in the bovine enzyme [5]. These observations suggest that the 40 kDa subunit may be less tightly bound to the enzyme. By contrast, the 30 kDa protein is seen prominently in essentially all purified vacuolar ATPase preparations, as would be expected if it is the homolog of the γ-subunit in F-ATPase.

The recent crystal structure of Abrahams et al. [26] for F1-ATPase shows clearly that each of the α- and β-subunits interacts directly with the γ-subunit. Furthermore, the key interacting regions of the γ-subunit are α helices. γ-subunits from different species are not highly conserved, but the regions of highest sequence identity are clustered in the first 50 and last 50 amino acids [27]. Like the γ-subunits, the *vma-4* gene products are most highly conserved near the C and N termini (Fig. 2), and their structures are predicted to be predominantly α-helical (Fig. 3). Taken together, the biochemical data and the sequence analyses suggest that in the vacuolar ATPase the *vma-4* gene product is a good candidate for the structural equivalent of the γ-subunit in the F-type ATPase. Very recently the sequence of a 28 kDa subunit of the vacuolar ATPase from yeast and bovine cells was reported [28]. The authors suggested this subunit could be the analog of the γ-subunit. However, this suggestion would not be consistent with biochemical data which suggest that the 28 kDa subunit is not essential for catalytic activity [23–25].

This work was supported by US Public Health Services grants GM28703 and GM 08123. We thank Karen Tenney for assistance in protein sequencing.

References

[1] Forgac, M. (1989) Physiol. Rev. 69, 765–796.
[2] Bowman, B.J., Vazquez-Laslop, N. and Bowman, E.J. (1992) J. Bioenerg. Biomembr. 24, 361–370.
[3] Zerez, C.R., Weiss, R.L., Franklin, C. and Bowman, B.J. (1986) J. Biol. Chem. 261, 8877–8882.
[4] Arai, H., Terres, G., Pink, S. and Forgac, M. (1988) J. Biol. Chem. 263, 8796–8802.
[5] Bowman, B.J., Dschida, W.J., Harris, T. and Bowman, E.J. (1989) J. Biol. Chem. 264, 15606–15612.
[6] Gogarten, J.P., Starke, T., Kibak, H., Fishman, J. and Taiz, L. (1992) J. Exp. Biol. 172, 137–147.
[7] Mandel, M., Moriyama, Y., Hulmes, J.D., Pan, Y.C., Nelson, H. and Nelson, N. (1988) Proc. Natl. Acad. Sci. USA 85, 5521–5524.
[8] Kane, P.M. and Stevens, T.H. (1992) J. Bioenerg. Biomembr. 24, 383–393.
[9] Mandala, S. and Taiz, L. (1986) J. Biol. Chem. 261, 12850–12855.
[10] Sze, H., Ward, J.M. and Lai, S. (1992) J. Bioenerg. Biomembr. 24, 371–381.
[11] Wieczorek, H. (1992) J. Exp. Biol. 172, 335–343.
[12] Stone, K., LoPresti, M., Crawford, J., DeAngelis, R. and Williams, K. (1989) in A Practical Guide to Protein and Peptide Purification for Microsequencing (Matsudaira, P.T., ed.), pp. 31–47 , Academic Press, San Diego, CA.
[13] Foury, F. (1990) J. Biol. Chem. 265, 18554–18560.
[14] Vollmer, S. and Yanofsky, C. (1986) Proc. Natl. Acad. Sci. USA 83, 4869–4873.
[15] Metzenberg, R. and Grotelueschen, J. (1992) Fungal Genet. Newslett. 39, 50–58.
[16] Bowman, B.J., Dschida, W.J. and Bowman, E.J. (1992) J. Exp. Biol. 172, 57–66.
[17] Gurr, S.J., Unkles, S.E. and Kinghorn, J.R. (1987) in Gene Structure in Eukaryotic Microbes (Kinghorn, J.R., ed.), pp. 93–139 , IRL Press, Oxford.
[18] Hirsch, S., Strauss, A., Masood, K., Lee, S., Sukhatme, V. and Gluck, S. (1988) Proc. Natl. Acad. Sci. USA 85, 3004–3008.
[19] Graf, R., Harvey, W.R. and Wieczorek, H. (1994) Biochim. Biophys. Acta 1190, 193–196.
[20] Ho, M.N., Hill, K.J., Lindorfer, M.A. and Stevens, T.H. (1993) J. Biol. Chem. 268, 221–227.
[21] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) J. Mol. Biol. 120, 97–120.
[22] Senior, A.E. (1990) Annu. Rev. Biophys. Chem. 19, 7–41.
[23] Peng, S.B., Stone, D.K. and Xie, X.S. (1993) J. Biol. Chem. 268, 23519–23523.
[24] Xie, X.S. and Stone, D.K. (1988) J. Biol. Chem. 263, 9859–9867.
[25] Puopolo, K., Sczekan, M., Magner, R. and Forgac, M. (1992) J. Biol. Chem. 267, 5171–5176.
[26] Abrahams, J.P., Leslie, A.G., Lutter, R. and Walker, J.E. (1994) Nature 370, 621–628.
[27] 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. (1985) J. Mol. Biol. 184, 677–701.
[28] Nelson, H., Mandiyan, S. and Nelson, N. (1995) Proc. Natl. Acad. Sci. USA 92, 497–501.
[29] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[30] Chou, P.Y. and Fasman, G.D. (1978) Annu. Rev. Biochem. 47, 251–276.