ELSEVIER

Biochimica et Biophysica Acta 1273 (1996) 77–83

BBA Biochimica et Biophysica Acta

Short sequence-paper

Isolation of the *vma-6* gene encoding a 41 kDa subunit of the *Neurospora crassa* vacuolar ATPase, and an adjoining gene encoding a ribosome-associated protein

Vladimir I. Melnik [1], Barry J. Bowman *

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

Received 27 September 1995; accepted 7 November 1995

Abstract

The *vma-6* gene, encoding a membrane-associated subunit of the vacuolar H^+-ATPase from *Neurospora crassa*, was cloned and sequenced. The gene contains three small introns and encodes a protein of 41 005 Da. When compared with homologous polypeptides from other species, the *N. crassa* protein contains a unique glycine-rich region. Three conserved cysteine residues, previously unrecognized, have been identified. An unrelated gene encoding a protein of 31 701 Da was found 2.1 kb downstream of *vma-6*. The second gene appears to encode the *N. crassa* homolog of a ribosome-associated protein identified previously in several plant and mammalian cells, and was named *rap-1*.

Keywords: ATPase, vacuolar; Vacuole; Proton pump; Ribosome; (*N. crassa*)

Vacuolar, or V-type H^+-ATPases are multisubunit proton pumps that acidify intracellular compartments in eukaryotic cells and provide the driving force for the secondary transport of various solutes [1]. In *Neurospora crassa*, the vacuolar ATPase has a molecular mass of approx. 700 kDa and is composed of at least nine different subunits [2]. The components of the enzyme are organized into two sectors. A hydrophilic sector named V_1 protrudes from the membrane and is the site of ATP hydrolysis. A hydrophobic sector, named V_0, forms the proton conducting pathway through the membrane. V_1 and V_0 sectors are separable in vitro, and recent data from two very different organisms, the yeast *Saccharomyces cerevisiae* [3] and the insect *Manduca sexta* [4], suggest that reversible dissociation of V_1 and V_0 may regulate proton pumping in vivo. Furthermore, we previously reported that oxidation of the enzyme from *N. crassa* can cause the dissociation of V_1 and V_0 sectors [5].

The V_0 sector contains a polypeptide described as the 39 kDa subunit in bovine cells [6] and the 36 kDa subunit in yeast cells [7]. The gene encoding this subunit has been isolated from only a few organisms, and some discrepancies were reported in determination of the beginning of the protein coding region [7]. Analysis of the sequences has not revealed typical membrane-spanning regions, even though the subunit co-purifies with the membrane sector. We have recently isolated a gene from *N. crassa* which encodes a 41 kDa polypeptide of the V_0 sector. Although slightly different, this polypeptide appears to be homologous to the 36–39 kDa subunit described in other organisms [6–8].

Isolation of the N. crassa vma-6 gene. A 0.83 kb fragment of *N. crassa* genomic DNA was amplified by the polymerase chain reaction using degenerate primers corresponding to conserved regions of the bovine and yeast genes [6,7]. Their positions are shown overlined in Fig. 1. The fragment was about 100 bp larger than predicted, suggesting the presence of an intron, which was subsequently confirmed. The 0.83 kb fragment was used to screen a cDNA library in λZAP (obtained from M. Sachs, Oregon Graduate Institute). Two clones of 0.4 kb and 1.5 kb were identified. Both were sequenced; the smaller clone (0.4 kb) was contained entirely within the larger (1.5 kb) polyadenylated clone.

* Corresponding author. Fax: +1 408 4593139; e-mail: bowman@biology.ucsc.edu.

[1] Present address: Renal Division, Box 8126 Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110, USA.

SSDI 0005-2728(95)00144-1

The larger cDNA was used to screen an *N. crassa* genomic library in the vector pMOcosX (Fungal Genetics Stock Center, University of Kansas, Kansas City, KS). Two overlapping clones were identified, #G11-2A and #X2-8B. The complete *vma-6* gene was located in #X2-8B, while #G11-2A contained only the region of the gene coding for the last 20 amino acids (starting at nt #2623 in Fig. 1). A 7.5 kb *Eco*RI fragment of #X2-8B was cut into smaller fragments which were subcloned into an M13 phage vector. Both cDNA clones and the first 6119 nucleotides of the 7.5 kb *Eco*RI genomic fragment were sequenced. Fig. 1 shows sequence of genomic DNA, the region corresponding to the cDNA for *vma-6*, and the downstream open reading frame which encodes the *rap-1* gene. To confirm that the *vma-6* gene encoded a subunit of the ATPase, we expressed the gene in *E. coli* and raised a rabbit polyclonal antibody. This antibody recognized a 40 kDa polypeptide in purified vacuolar membranes from *N. crassa* (Steinhardt, A.S. and Bowman, B.J., unpublished data).

Chromosomal mapping of the vma-6 gene. Chromosomal location was determined by analysis of restriction fragment length polymorphisms [9] using DNA digested with *Eco*RI. The *vma-6* gene was located near the centromere of linkage group II. Two other V-ATPase genes, *vma-2* and *vph-1*, map to this chromosome but are separated by at least 10 map units on either side [10] (Bowman, E.J., personal communication).

Characteristics of the vma-6 gene and its product. The *vma-6* gene contains three small introns, of 56, 76 and 60

```
   1  GATCTTGGCCAACTCGGGGAAGTGCCAGCCGTACCATTCCTTGACACGCATGGCGTAGGTGTTGAGCTCC
  71  TTGTCAAGCTCGTCGAGGAGCGAGACGGCGTGGACGATCATGACATCGACCTTCTCGGGGGAGAACTTGA
 141  GCTTGTGGCGCGAGAGAGAGTGGGAGAGACCGAGGGACATCTCCTTGAAGTTCTCGGGGAGCATGCCAGG
 211  AATGAGCTCGGGAAGGTACTGGCGGATGGCACGGAAGAGGTCGGTGGTGGTGGAGTCGGCAATGGGCTGG
 281  ATGTCGAGGTTGGGGATCTTGTTGATGGCAGCCGGAGCTTGCTCTCGGCAACGGCGAGGGTGACCTTCTT
 351  CTCATTTCCAATCTCGGCAAGCAGGCTGTTCAGCTTGGGAGTGACCTTGCCCTCAACAACACCAGCGATT
 421  TCCTCGAGGGCAGTGGCGGCGCTCTCGAACTTGGCGAACTCCTTGTACTTGATCCTGTAGGCACACGCGT
 491  GTTGGTCAATTGCATCTTGTTGCCATGATGTGACGAGGCAGCTTACTCCTTGGTGATCTTTTCAACGCTG
 561  TTGAGACGCTCGGCGAGGTTGTCCGAGGACAGCAGCTTCTTGTCTGCCGCCTTGAACAGGCCATAACTGT
 631  AGGGACTTGTTAGACATGTTGGACTATTATGGTAGGTAGGTGTTGTTGAGGTTCGAGACGGAAGGGGAGA
 701  CCTCTCGGTCTGATCGCAGCCGAATTGGACGATTCAAACTTGAAGCCTCCGCATATGCAGGCATGGACGA
 771  CGCCGCCGCCGTTGTGTATGGACGTGAATTTGATGTGATGGCACTTACCCGGCACTGGTTTCTGTTAGGA
 841  TAAAGAGAGGCATGGTTGCGGTTTGGTGTTTGGATCCCGAACGCGCTGGCAATTGTCGTTGGGAAACAGC
 911  AAGTGGGTGTGCTGTTTTTTTGCGATATTCACCACGAGCCTCGAGGGGGAGAAAAAAGTCGAGAGCAAAG
 981  AACAATTTTTTTCCACCAGACTTTGGAGCACCCACTTCTGCCCACACTCTGCCTGCCCACTCACCGAAAT
1051  CCATGGCGCCCGTTCCCCCAATCAGAGCCCGGTGTGGGGTCGAAGAGGCCAGAGGTGGCCCGAGCTTGGC
1121  TATTGCACGTATCTTTTCAACGGCCAACTTTGTCTCCATCAACCCTCGCTCTGGCCGCTCTGAATTGGAG
1191  ATGCCAAGACCCGTCATGCACGGTCCAACCCTCTCTCTGAAGCGCCGATGGTGGGCGCCCTGTCGCAGCC
1261  TTCCTGCGCCTGTGACCTCGATGGACAACTGGAGCTCCAACAACGACAAGAGCAGCCCAGACAGGGAGTG
1331  CCAGCCATCTGTTGTTGTTGAACACTCGCTCGCTTGAACATATCACCACCTCCAACTCCAACCGCAGACA
1401  TGGAGGGCCTATTATTCAACGTCAACAATGGgtatgttgtctcatttctaatttgcagatacagtgctga
        M  E  G  L  L  F  N  V  N  N  G

1471  ccatcattcatcaacagCTACATTGAGGGTATCGTCCGTGGATACAGGAACAGCCTGTTGACGAGCACCA
                          Y  I  E  G  I  V  R  G  Y  R  N  S  L  L  T  S  T

1541  ACTATACCAACATGACGCAATGCGAGTCGATTGATGgtgagcctctctctcccatctccatccccgaccc
        N  Y  T  N  M  T  Q  C  E  S  I  D

1611  atttcacctccatccacatcgccctaacaacattgcccacagACCTTAAACTCCAACTCGGACCTGCATA
                                              D  L  K  L  Q  L  G  P  A  Y

1681  CGGCGACTTCCTCGCCTCTCTCCCGCCCAAGCCTTCGACCTCCGCTCTCGCTGCCAAGACCACCGACAAG
        G  D  F  L  A  S  L  P  P  K  P  S  T  S  A  L  A  A  K  T  T  D  K

1751  CTGGTCTCCGAGTTCCGCTATGTCCGCGCCAACGCGGCCGGCTCCCTGGCCAAGTTCATGGACTACCTCA
        L  V  S  E  F  R  Y  V  R  A  N  A  A  G  S  L  A  K  F  M  D  Y  L

1821  CCTACGGCTACATGATCGACAACGTCGCCCTGCTCATCACGGGCACCCTTCACGAGCGCGACACGCGCGA
        T  Y  G  Y  M  I  D  N  V  A  L  L  I  T  G  T  L  H  E  R  D  T  R  E
```

Fig. 1. Sequence of the *vma-6* and *rap-1* genes. Nucleotides 1400–2686 and 4780–5945 represent the coding regions of *vma-6* and *rap-1* respectively. Positions corresponding to the degenerate primers used to amplify the *vma-6* gene are *overlined.* Introns are shown in *lower case* within coding regions. Two pairs of putative regulatory elements are shown in lower-case upstream of the *rap-1* gene. The complete nucleotide sequence is deposited in GenBank under accession number U36470.

bp, respectively. Two of the introns were found within the first 20% of the protein coding region as is typical for *N. crassa* genes [11]. The introns do not appear to occur at the boundaries of obvious domains in the protein. A polyadenylation site was found 352 bp downstream of the stop codon (Fig. 1).

The protein encoded by *vma-6* is 41 005 Da in size, made up of 364 amino acid residues with a calculated p*I* of 4.65. The protein is hydrophilic overall with no obvious membrane-spanning domains, even though it is known to behave as a membrane-associated subunit [2,6–8,12].

Alignment of the *N. crassa* protein sequence with other *vma-6* proteins (Fig. 2) shows that it is 46, 50 and 52% identical with the *S. cerevisiae, Dictyostelium discoideum*, and bovine counterparts, respectively, with an overall identity between the four species of 31.9%. An updated version of the bovine sequence was used for the alignment (GenBank accession number J04204). A human protein sequence deduced from cDNA [13] was not included in the alignment because it was truncated at the same position as the earlier version of the bovine protein [6]. Examination of the 5′ region of the human cDNA showed a possible protein coding region 90% identical to the revised bovine sequence but interrupted by several frameshifts (data not shown). It seems likely that in the reported human sequence the beginning of the protein coding region was not correctly identified.

Compared with other *vma-6* gene products, the *N. crassa* protein is larger by 8–19 amino acid residues mostly due to the presence of a unique glycine-rich stretch of 19 residues at positions 283–301 (Fig. 2). In *S. cerevisiae* this region is absent, and only 5–6 residues are observed in the other two species shown. These data suggest that this part of the protein is on the surface, or

```
1891  GCTGCTCGAGCGCTGCCACCCGCTCGGCTGGTTCGAGACCATGCCCGTCCTCTGCGTGGCCACCAACATT
        L  L  E  R  C  H  P  L  G  W  F  E  T  M  P  V  L  C  V  A  T  N  I

1961  GAGGAGCTGTACAACAGCGTCATGATCGAGACCCCACTGGCCCCGTACTTTAAGAGCAGCCTGTCGCTCC
        E  E  L  Y  N  S  V  M  I  E  T  P  L  A  P  Y  F  K  S  S  L  S  L

2031  AGGACCTGGACGAGCTCAACATCGAGATTGTGCGCAACACGCTTTACAAGAACTACCTCGAGGACTTTTA
        Q  D  L  D  E  L  N  I  E  I  V  R  N  T  L  Y  K  N  Y  L  E  D  F  Y

2101  CCACTTCGTCAACACCCACCCCGATATGGCCGGGACCCCGACTGCCGAGGTGATGTCGGAGTTGCTGGAG
        H  F  V  N  T  H  P  D  M  A  G  T  P  T  A  E  V  M  S  E  L  L  E

2171  TTTGAGGCGGATAGGCGGGCTATCAACATCACTCTTAACTCGTTCGGCACCGAGCTGTCCAAGGCGGACC
        F  E  A  D  R  R  A  I  N  I  T  L  N  S  F  G  T  E  L  S  K  A  D

2241  GTAAGAAGCTCTATCCTAACTTTGGGCAGCTGTACCCGGAGGGGACGCTCATGTTGTCGCGGGCAGATGA
        R  K  K  L  Y  P  N  F  G  Q  L  Y  P  E  G  T  L  M  L  S  R  A  D  D

2311  CTTTGAGGGTGTGAGGCTCGCTGTTGAGGGTGTGGCTGATTACAAGTCGTTCTTCGATGCTGCTGGCCTC
        F  E  G  V  R  L  A  V  E  G  V  A  D  Y  K  S  F  F  D  A  A  G  L

2381  GGAGGCGGTCCCAGCGGTCCCGGTAACATGGGCGGTGGCGGAACTGAGGGCAAGAGTCTGGAGGACATGT
        G  G  G  P  S  G  P  G  N  M  G  G  G  G  T  E  G  K  S  L  E  D  M

2451  TCTACCAGAAGGAGATGGAGATCTCCAAGATGGCGTTCACGAGGCAGTTCACCTATGCCATAGTGTATGC
        F  Y  Q  K  E  M  E  I  S  K  M  A  F  T  R  Q  F  T  Y  A  I  V  Y  A

2521  GTGGGTCAAGTTGAGAGAGCAGgtgagattagggttcccctgtgttgttcgttgagcatctgttgactaa
        W  V  K  L  R  E  Q

                                ──────────────────────────
2591  tgtgacatttagGAAATCCGCAATATCACTTGGATCGCCGAGTGCATAGCGCAGAACCAGAAGGAGAGGA
                    E  I  R  N  I  T  W  I  A  E  C  I  A  Q  N  Q  K  E  R

2661  TCAACAACTACATCAGCGTTTTCTAAAGAGCAGTAAGTTGTCGTAATGAGGTTTGATTTTCAGGTTGCAT
      I  N  N  Y  I  S  V  F  *

2731  TTTTATCAGAGGAACCTCGTCCTCCCCCTATGTTTTGCTTGAGCCGGGTTGTCAGTGCATGTTATCAATG

2801  ACCAAGAAATGGGAAGGGGAGAAAAGGCGTATTGGGCGCGTTACGGCCAAAGTATTATATCTGGACTCTT

2871  GGGGAACACAAAAGGGGTATCATCATCATCGACATGGCCTGCCCCCGCGTGCGCCCCTTGGTTTTCTCCG

2941  GTTATTGTTTATTGTTGTATTGTACATAGAGAAGGCCTCTTTGTTTTGTTACACACGCTTTCATGCTACG
                                  ↓ site of poly(A) addition

3011  ACTGGAAATCACATCTTCTTCCGCTCACTCACACACAAACACACACGCGACAACCGAACCGGGTGCGTTG

3081  AATGTATTAGAGGGTGATCTGACCCTTCCAGGTGGCGGGCGTCAGAGCAGGGTCTCGAGCATGTGATGCC

3151  ACCTTCCACTGCCTCTCTCCTTGCGCTCCATGTACCTCAGCAGTCTTTGGCGCTTGTGGCACAGCTCCGC

3221  AAACTGCGCTTGTTGTTCTTATCTCTGTTGCCGCCGTGGCCCTCGAGCGCCTTGCTGAGCGCCCGGATCT

3291  TGGACGTCAGGATGGCGATCTGCACCTCACTGCTGCCCGTGTCCGGGCCTGCCGCACCGGCTGGGGCTTG

3361  GGCTCCACTCCGCGCGCCAGCGGCGGCGTCGGCCTCTCGAGCGTCTGATCTGTAATGTGGCGACCGAAAG

3431  TCTCGATGCAGCGGCGGATGTTGGCGTGTTTGCGGTCTTTAGCACCACCATGCTCAAGCTTCGCAGTACG
```

Fig. 1 (continued).

acts as a flexible linker between subdomains within the polypeptide.

A previously unrecognized feature of the eukaryotic *vma-6* proteins is the presence of three conserved cysteine residues found at positions 36, 125 and 348 in *N. crassa* (Fig. 2). Data from several laboratories indicate that the vacuolar ATPase may be regulated by oxidation and reduction of cysteine residues [14–16]. Although most previous work has focused on the *vma*-1 gene product, a 67 kDa protein that contains the ATP-binding site, it is very possible that other subunits, such as the *vma-6* gene product, are involved [5].

Identification of the rap-1 gene. Preliminary sequence data from the *Eco*RI genomic fragment which contained *vma-6* indicated that a second open reading frame was present. Because little information is available regarding intergenomic regions in *N. crassa* we sequenced most of the *Eco*RI fragment. On the basis of sequence similarity the second gene appeared to encode the *N. crassa* homologue of a protein named p40 in yeast, mammals [18,19], *Drosophila* [20] and *Arabidopsis* [21]. Originally reported to be a precursor for the extracellular 67 kDa laminin receptor which is highly expressed in some tumor cell lines [19,22], the protein was subsequently found to be intracellular, associated with ribosomes [21,23] and/or cytoskeleton [24]. In *S. cerevisiae* the *NAB1A* and *NAB1B* genes encode homologues of the p40 protein which are 90% identical to each other [17] (Ellis, S.R., University of Louisville, personal communication). Disruption of either gene caused a reduction in growth rate and a shift from larger to smaller polysomes. Disruption of both genes was lethal. We named the *N. crassa* gene *rap-1* (*r*ibosome *a*ssociated *p*rotein).

The *rap-1* gene appears to contain four introns, the first

```
3501  CTCAAGGGCGGCGACGGCTTGGCGTGTCTGGCGGCGTGCTCGAGCTCCTCGTCCTCCTTCTCCGGGTCGC
3571  CGTCAAGCGGTCCTCGGCCTTGATGGGCTTCAAGATGTGCTCGGCGTACTGAATGGCGCTGTCGAACTCG
3641  TCCTTGTTGAGCAGGTAATTCCGTAGTGCGGCGAGGTTGGGAGGGGCTTGGGTTCCTCAACAAGCTGGCC
3711  GTACGGGGCGACCTTGGGCAGAGACGGAAGCCGCCGGCGCTGTCGAAGGATTCGCAGAAGGGGGTCGT
3781  GATACCCTTGACCGGGTCGCCCATGCGCTCGCTTGGCTGAGCTCGGCCTGGCGCTTCACGTTGACGGCCT
3851  TGCGCTGTTGCGCCTGGGCCCAGCCGTAAGGTCCTGCTTCATTTTACGCTTCTTTTCTTTCTGGGAGATG
3921  TTGGCGGTTTGGATGAGAGGTTGCAGGGCCTGCGCCGGAGAGGCCACGGCGGGGCGAAGACATACTGTCG
3991  TTTCCAGAGCAACATGTCAGTACGTAGGTGTGTTTGGCGGGGGACATCATGGACGGCATGGCAAAAAGTT
4061  GAAAGGTAAATAAAAGTCCACGTACGCGTGAGGCTGCGAACGCCTTGAGGGCCAGGTAACCTGGGAGGCA
4131  TGGTTGGTTGGTTGGGTGTGTGGTGGTGGTGGAGGTGCACCGGCGTGGTCGGTCTCGAAGTGCAAGT
4201  CGGCGCAACAACTTTCTTTCAAGTCTCGCAGCTTggcccggcgGTCCCTCCTCGTCcctccagGTCACGG
4271  TCCGGCCGATGTTGAACAACTTCCCGTGTCGTCCGCACCTGAACCAAAATTTGGGTTCCAAGTTCGCTAG
4341  GCTAGAACAGGGTTAGTCCATCTCTCAACTCGAGCTTGCAGACTGCGAAGTGGATTCCACTTCACTGATC
4411  ACCTGACGTCGGCCACCAGGTGAAGCAACTCAAATTCAGAAAGAGGCCTGTCGAGGTTACGATTTAATGG
4481  GGCTAATCGTATTTGAAGCTATTGATTGGATGCAGTCTGATGTCCAAGGAGAAGAATACGTACCATGCTG
4551  CAATGGCTTGGCACGGTCAACTGTCAACCAACtgcacagccCTGGCGGCGAGTCGGCGCActccggCACT
4621  CCACAAAATTGAATTTGTGTGCGAAAATCAGAGGCAGAAAGCTCGGACGTCCCGCGACTTTGGCTGGCTG
4691  GTTGAATTAGAAGAAAATATATTCACCCTCCACGCCGTCTTCCTGACACCTCCTCAGTCCTCACACCGCA
4761  AGCCAGAGATCACGTCAAGATGGCGCCCGCTAACCTCCCCTCCATCTTCAATGCCACGAGCACCGATATT
                          M  A  P  A  N  L  P  S  I  F  N  A  T  S  T  D  I
4831  GAGCAGCTCCTTGCTGCTCAGTGCCACATCGGCTCCAAGAAgtgcgtttttgcgaaatccacatcgccgg
       E  Q  L  L  A  A  Q  C  H  I  G  S  K  N
4901  cgacatggctggagaaaatgagacctcgaaatgctgactttgtgcgtttgcgaatagCCTCGGTGTTCAC
                                                              L  G  V  H
4971  ATGCAGCgtatcactccctatcgatatataacacttgtcgaagctacaacgacttgcgatcgcacggaca
       M  Q
5041  aatttctgacctgacttgactctgcagCTTACCTCTGGAAGACTCGCGCCGACGGTGTCAACATCTTGAA
                                  P  Y  L  W  K  T  R  A  D  G  V  N  I  L  N
5111  CGTTGGCAAGACCTGgtatgtggaggatcgatcgatcgcatgagtggatggtgtccaagtggaggagaca
        V  G  K  T  W
5181  tggtcaccaggatacctcgaacaaaggccacaatggctgaacgtgctggctcaccacagGGAGAAGATCG
                                                                  E  K  I
5251  TCCTTGCTGCCCGCATCATCGCCGCCATCGACAACCCGGCCGATATCTGCGTTATCTCTGCCCGTCCAAT
      V  L  A  A  R  I  I  A  A  I  D  N  P  A  D  I  C  V  I  S  A  R  P  I
```

Fig. 1 (continued).

```
5321  CGGTCAGCGTGCCGTCCTGAAGTTCGCCGCCCACACCGGTGCTCAGGCCATCGCCGGTCGCTTCACCCCC
        G  Q  R  A  V  L  K  F  A  A  H  T  G  A  Q  A  I  A  G  R  F  T  P

5391  GGTTCTTTCACCAACTACATCACCCGCTCTTTCAAGGAGCCCCGCCTCATCATCGTCACCGACCCCCGCA
        G  S  F  T  N  Y  I  T  R  S  F  K  E  P  R  L  I  I  V  T  D  P  R

5461  CTGATGCCCAGGCCATCAAGGAGGCTTCCTACGTCAACATCCCCGTCATTGCCCTCTGCGATGCCGACTC
        T  D  A  Q  A  I  K  E  A  S  Y  V  N  I  P  V  I  A  L  C  D  A  D  S

5531  GCCCACCGAGTACGTCGATGTTGCCATCCCCACCAACAACAAGGGTCGCCACGCCATCGCGTGCGTCTGG
        P  T  E  Y  V  D  V  A  I  P  T  N  N  K  G  R  H  A  I  A  C  V  W

5601  TGGATGCTTGCCCGTGAGGTCCTCCGCCTCCGCGGTACCATCTACAACCGCGAGACCCCCTGGGACGTCA
        W  M  L  A  R  E  V  L  R  L  R  G  T  I  Y  N  R  E  T  P  W  D  V

5671  TGGTCGATCTTTACTTCTgtatgtatccccggagcccccgtgaccgtctgagggaccatgtctctaacc
        M  V  D  L  Y  F

5741  attaacctacagACCGCGACCCCGAGGCTGAGGCCGAGGAGAAGGTTGAGGAGGAGAAGCTCCCTGGCGT
                    Y  R  D  P  E  A  E  A  E  E  K  V  E  E  E  K  L  P  G  V

5811  TGAGGAGGAGGGTGTTGCCGCCATTGAGTCCGGCTTCCCCGCCACTGGTGACTGGGAGGCTGCCCCCGCT
        E  E  E  G  V  A  A  I  E  S  G  F  P  A  T  G  D  W  E  A  A  P  A

5881  GGCTTCCCCGCCACCGGTGAGTGGTCCGATGCCGCTCCCGGTGCCGCTGCCCCCAACTGGGACGCTACTG
        G  F  P  A  T  G  E  W  S  D  A  A  P  G  A  A  A  P  N  W  D  A  T

5951  CCCCTGCCACCACTGCTGACTGGGCCGCCACCGAGGCAAGGAGTCCAGCTGGTAAACAGTTGTAATCGAA
        A  P  A  T  T  A  D  W  A  A  T  E  A  R  S  P  A  G  K  Q  L  *

6021  AAAAACACATTCGGTTTCTGGGTCAAAAATGGCGTAATGGAGGCGTGATGGAGACTGGACTTGTGGCTTG

6091  AACCTCGCCACCGATTTCCAACACGAATT
```

Fig. 1 (continued).

three located within the initial 20% of the coding region (see Fig. 1). It encodes a protein of 31 701 Da composed of 293 amino acids with an estimated p*I* of 4.8. Alignment of *N. crassa* p40 with the yeast, *Drosophila* and human homologues is shown in Fig. 3. Overall, 40% of the residues are identical in all four sequences. The C-terminal

```
                                                                    ↓
nspora    1            MEGLLPNVNNGYIEGIVRQYRNSLLTSTNYTNMTQCESIDDLKLQL·GPAY
yeast     1            MEGVYPNIDNGFIEGVVRQYRNGLLSNNQYITLTQCDTLEDLKLQLSSTDY
dictyo    1  MGLFGGRKHGGLFTPNKDDQYLEAILRQFKKGILSRADYNNLCQCDNLEDMKMHFISTDY
bovine    1        MSFFPELYPNVDNGYLEGLVRQLKAGVLSQADYLNLVQCETLEDLKLHLQSTDY

nspora   51  QDFLASLPPKPST·SALAAKTTDKLVSKFRYVRANAAGSLAKFMDYLTYGYMIDMVALLI
yeast    52  QNFLSSVSSESLTTSLIQEYASSKLYHKFNYIRDQSSGSTRKFMDYITYGYMIDMVALMI
dictyo   61  QDFLAGEPSPIHTTT·IAEKATGKLVSKFNHIRNQAVEPLSTFMDFISYGYMIDMVLLI
bovine   55  QNFLANEASPL·TVSVIDDRLKEKMVVKFRHMRNHAYEPLASFLDFITYSYMIDMVILLI
                                  ↓
nspora  110  TGTLHERDTREKLLERCHPLGWIETMPVLCVATNIEELYNSVMIETPLAPYFKSSLSL·QD
yeast   112  TGTIRDRDKGEILQRCHPLGWFDTLPTLSVATDLESLYETVLVDTPLAPYFKNCFDTAEE
dictyo  120  TGTLHERDISELVDKCHPLGLFKSMATLSVVHNVADLYNNVLIDTPLAPYIQGCLS·EED
bovine  114  TGTLHQRSIAELVPKCHPLGSFEQMEAVNIAQTPAELYNAILVDSP·GGFFPG·LHLEQD

nspora  169  LDELMKEIVRNTLYKNYLEDFYHFVNTHPDMAGTPTAEVMSELLEFEADRRAINITLNSF
yeast   172  LDDMMKEIIRNKLYKAYLLDFYNFV·T·EEIPE·PAKECMQTLLGFEADRRSINIALMSL
dictyo  179  LDEMMKEIIRNTLYKAYLEDFYNY·C·KY·LGGQ·TELIMSDILKFEADRRSINITINSF
bovine  172  LDEMMKEIIRNTLYKAYLESFYKF·CT·L·LGGT·TADAMCPILEFEADRRAFITINSF

nspora  229  G·TELSKAD·RKKLYPNFSQLYPEGTLMLSRADDFEGVRLAVEGVADYKSFFDAAGLGGG
yeast   229  QSSDIDP·DLKSDLLPNIGKLYPLATFHLAQADDFEGVRAALANVYEYRG·FLETG····
dictyo  235  GATELSKDD·REKLYPSLGLLYPEGTSKLGKAEDVDQVRGILEVYSTYRNFFSD·GVNN·
bovine  228  G·TELSKED·RAKLFPHCGRLYPEGLAQLARADDYEQVKNVADYYPEYKLLFEGAG····

nspora  287  PSGPGNMGGGGTEGKSLEDMFYQKEMEISKMAFTRQFTYAIVYAWVKLREQEIRNITWLA
yeast   283  ··············NLEDHFYQLEMELCRDAFTQQFAISTVWAWMKSKEQEVRNITWLA
dictyo  292  ············E·KSLEDSFFEHEVHLNRMAFEDQYGYGVFYAYIKLREQEIRNIVWLA
bovine  282  ·SNPG·····D·KTLEDRFFEHEVKLNKLAFLNQFHFGVFYAFVKLKEQECRNIVWLA
              ↓
nspora  347  ECIAQNQKERINNYISVF*
yeast   328  ECIAQNQRERINNYISVY*
dictyo  339  ECISQNMKQKMNQYIPIF*
bovine  333  ECIAQRHRAKIDNYIPIF*
```

Fig. 2. Alignment of the vma-6 gene product with three homologous proteins. Identical residues are shown in *bold* and *underlined*. Gaps are indicated by *dots*. Sequences are shown in the order: *N. crassa* (this work), *S. cerevisiae* [17], *D. discoideum* [8] and *B. taurus*, (updated version, GenBank No. J04204) [6]. Conserved cysteine residues are indicated by arrows.

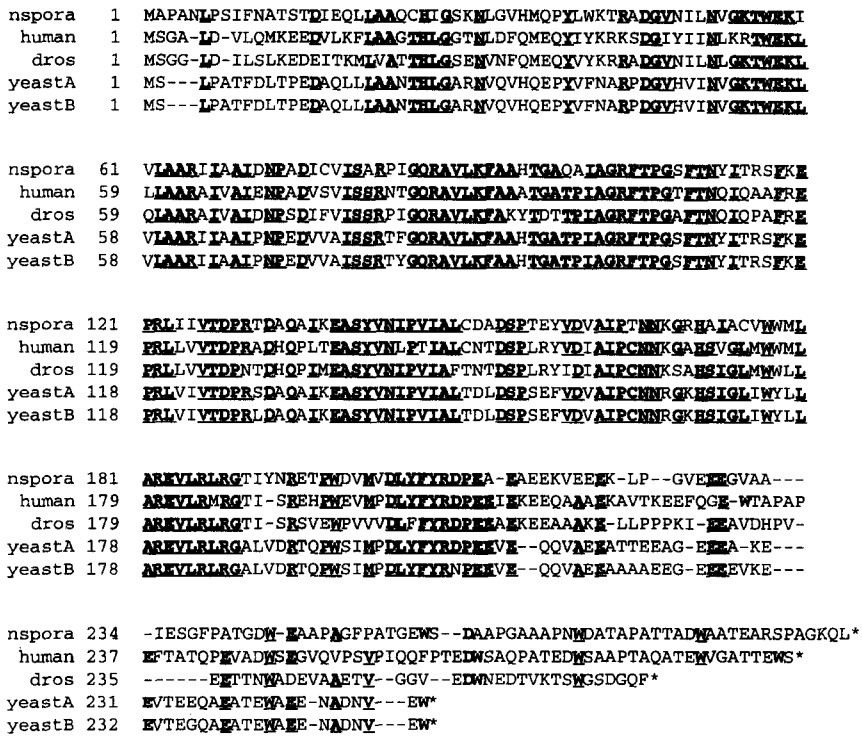

Fig. 3. Sequence comparison of the p40 proteins. Identical residues are shown in *bold* and *underlined*. Gaps are indicated by *dashes*. Proteins are shown in the order: *N. crassa* (this work), human [19], *D. melanogaster* [20] *cerevisiae* Nab1A [17] and Nab1B (Ellis, S.R., University of Louisville, personal communication).

half is rich in acidic amino acids and has an interesting pattern of evenly spaced tryptophan residues.

If *rap-1* encodes a ribosome associated protein, it may contain upstream elements common to other *N. crassa* genes which encode ribosomal proteins. Carbon and growth regulation of ribosomal genes appears to be controlled by two elements, spaced 12–16 bp apart (A/G, GC C/T N(N) A/G GCC/T and C C/T CC A/G C/G) (Brett Tyler, University of California, Davis, personal communication). The *rap-1* gene appears to have two of these elements, at −546 to −518 and −198 to −164 (numbering relative to the translation startpoint, shown in Fig. 1). The intergenic region is approximately 1200 bp, from the poly(A) site for the *vma*-6 gene to the first of the putative regulatory elements for *rap-1*. Both *vma*-6 and *rap-1* may be 'housekeeping genes' which encode fairly abundant products expressed under most growth conditions [25,26]. It will be interesting to see if other 'housekeeping genes' are encoded in the same region of the genome.

This work was supported by US Public Health Services grant GM28703.

References

[1] Forgac, M. (1989) Physiol. Rev. 69, 765–796.
[2] Bowman, B.J., VazquezLaslop, N. and Bowman, E.J. (1992) J. Bioenerg. Biomembr. 24, 361–370.
[3] Kane, P.M. (1995) J. Biol. Chem. 270, 17025–17032.
[4] Sumner, J.P., Dow, J.A., Earley, F.G., Klein, U., Jager, D. and Wieczorek, H. (1995) J. Biol. Chem. 270, 5649–5653.
[5] Dschida, W.J.A. and Bowman, B.J. (1995) J. Biol. Chem. 270, 1557–1563.
[6] Wang, S.Y., Moriyama, Y., Mandel, M., Hulmes, J.D., Pan, Y.C., Danho, W., Nelson, H. and Nelson, N. (1988) J. Biol. Chem. 263, 17638–17642.
[7] Bauerle, C., Ho, M.N., Lindorfer, M.A. and Stevens, T.H. (1993) J. Biol. Chem. 268, 12749–12757.
[8] Temesvari, L., Rodriguez-Paris, J., Bush, J., Steck, T.L. and Cardelli, J. (1994) J. Biol. Chem. 269, 25719–25727.
[9] Metzenberg, R. and Grotelueschen, J. (1992) Fungal Genetics Newsl. 39, 50–58.
[10] Bowman, B.J., Allen, R., Wechser, M.A. and Bowman, E.J. (1988) J. Biol. Chem. 263, 14002–14007.
[11] Bowman, B.J., Dschida, W.J. and Bowman, E.J. (1992) J. Exp. Biol. 172, 57–66.
[12] Zhang, J., Myers, M. and Forgac, M. (1992) J. Biol. Chem. 267, 9773–9778.
[13] van Hille, B., Vanek, M., Richener, H., Green, J.R. and Bilbe, G. (1993) Biochem. Biophys. Res. Commun. 197, 15–21.
[14] Feng, Y. and Forgac, M. (1994) J. Biol. Chem. 269, 13224–13230.
[15] Hager, A. and Lanz, C. (1989) Planta 180, 116–122.
[16] Taiz, L., Nelson, H., Maggert, K., Morgan, L., Yatabe, B., Taiz, S.L., Rubinstein, B. and Nelson, N. (1994) Biochim. Biophys. Acta 1194, 329–334.
[17] Davis, S.C., Tzagoloff, A. and Ellis, S.R. (1992) J. Biol. Chem. 267, 5508–5514.
[18] Rao, C.N., Castronovo, V., Schmitt, M.C., Wewer, U.M., Claysmith, A.P., Liotta, L.A. and Sobel, M.E. (1989) Biochemistry 28, 7476–7486.
[19] Yow, H.K., Wong, J.M., Chen, H.S., Lee, C.G., Davis, S., Steele, G.D., Jr. and Chen, L.B. (1988) Proc. Natl. Acad. Sci. USA 85, 6394–6398.

[20] Melnick, M.B., Noll, E. and Perrimon, N. (1993) Genetics 135, 553–564.
[21] Garcia-Hernandez, M., Davies, E. and Staswick, P.E. (1994) J. Biol. Chem. 269, 20744–20749.
[22] Wewer, U.M., Liotta, L.A., Jaye, M., Ricca, G.A., Drohan, W.N., Claysmith, A.P., Rao, C.N., Wirth, P., Coligan, J.E., Albrechtsen, R. and et al. (1986) Proc. Natl. Acad. Sci. USA 83, 7137–7141.
[23] Auth, D. and Brawerman, G. (1992) Proc. Natl. Acad. Sci. USA 89, 4368–4372.
[24] Keppel, E. and Schaller, H.C. (1991) J. Cell. Sci. 100, 789–797.
[25] Sachs, M.S. and Yanofsky, C. (1991) Dev. Biol. 148, 117–128.
[26] Shi, Y.G. and Tyler, B.M. (1991) Nucleic Acids Res. 19, 6511–6517.