YEAST VOL. 8: 787–790 (1992)

V : Yeast Sequencing Reports

AFG1, a New Member of the *SEC18*–NSF, *PAS1*, *CDC48*–VCP, TBP Family of ATPases

YANG-JA LEE AND REED B. WICKNER

Section on Genetics of Simple Eukaryotes, Laboratory of Biochemical Pharmacology, National Institute of Diabetes, Digestive and Kidney Disease, Bldg. 8, Room 207, National Institutes of Health, Bethesda, MD 20892, U.S.A.

Received 6 May 1992; accepted 11 May 1992

We have sequenced a gene that encodes a 377 amino acid putative protein with an ATPase motif typical of the protein family including SEC18p (NSF = N-ethyl maleimide-sensitive fusion protein; vesicle-mediated endoplasmic reticulum to Golgi protein transfer), PAS1p (peroxisome assembly), CDC48p (VCP = valosin-containing protein; cell cycle) and TBP1 (Tat-binding protein). This gene, *AFG1* for ATPase family gene, also has substantial homology to these proteins outside the ATPase domain. *AFG1* is located on chromosome V immediately centromere-proximal to *MAK10*.

KEY WORDS — *Saccharomyces cerevisiae*; ATPase; chromosome V.

INTRODUCTION

Recently a family of proteins has been described, some of which have demonstrated ATPase activity and all of which have an ATPase domain with the consensus pattern, h..hG....[KR]G-[ILV]LLYGPPG[TC]GKTL[ILM], and substantial homology to each other beyond this region.

Address for correspondence: Reed B. Wickner, Bldg. 8, Room 207, National Institutes of Health, Bethesda, MD 20892, U.S.A.

The yeast SEC18p and the highly homologous mammalian N-ethyl maleimide-sensitive fusion

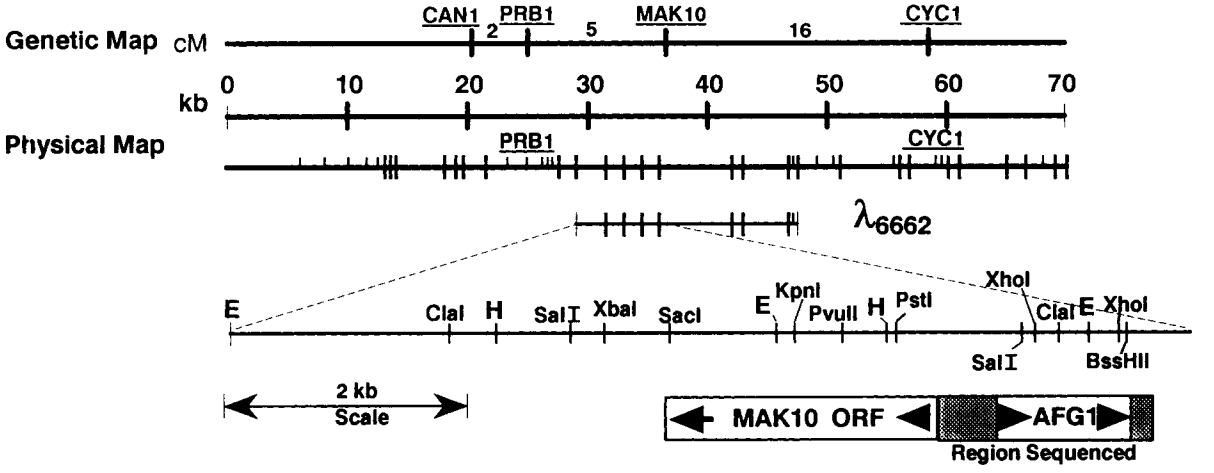

Figure 1. Location of the *AFG1* gene on the genetic and physical maps of a region of chromosome V. Genetic markers near *MAK10* on the left arm of chromosome V are shown with the distances between them in centimorgans. The physical map with locations of *CYC1* and *PRB1* is from Riles *et al.* (1992). Vertical lines passing through the horizontal axis are *Hin*dIII or *Eco*RI sites whose location is known, while vertical lines that are only above the horizontal axis represent sites whose order has not been determined and are shown in decreasing size. The detailed physical map of λ_{6662} and the location of the *MAK10* ORF is from Lee and Wickner (submitted). E = *Eco*RI, H = *Hin*dIII.

0749–503X/92/090787–04 $07.00

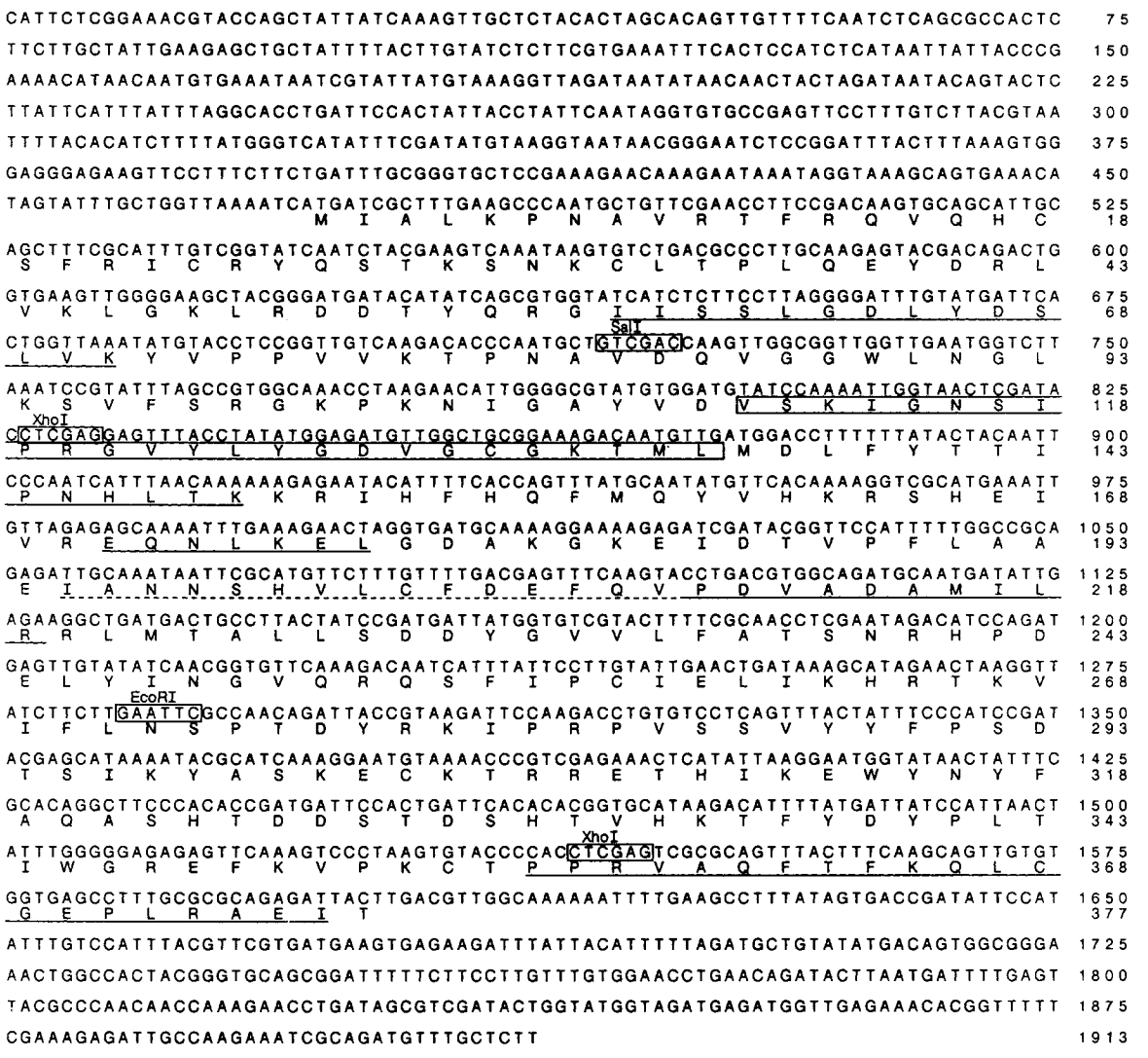

Figure 2. Sequence of the *AFG1* gene. Bases 1 to 3 are the complement of the ATG start codon of *MAK10*, which is encoded on the opposite strand (Lee and Wickner, submitted). Regions of homology with the SEC18–NSF, PAS1, CDC48–VCP and TBP1 proteins (see Figure 3) are underlined, and the special ATPase motif characteristic of these genes is boxed.

protein (NSF) are essential for transport of proteins from the endoplasmic reticulum to the Golgi (Novick *et al.*, 1981; Eakle *et al.*, 1988; Wilson *et al.*, 1989; Weidmann *et al.*, 1989; Kaiser and Schekman, 1990; Clary *et al.*, 1990). This transport takes place via small vesicles whose fusion with the Golgi requires (interchangeably) SEC18p or NSF. This activity requires ATP and another protein (yeast SEC17p or mammalian SNAP).

Yeast PAS1p is necessary for peroxisome biogenesis (Erdmann *et al.*, 1991). While it has the special ATPase motif of this group of proteins, its sub-cellular location and precise role are unknown.

CDC48p of *S. cerevisiae* is required for progression of the cell cycle (Moir *et al.*, 1982). The *cdc48* mutants arrest as cells with large buds having abnormal microtubule structures extending throughout the cytoplasm. CDC48p is loosely associated with particles in the microsomal fraction. The sequence of CDC48p (Frohlich *et al.*, 1991) is 70% identical to that of valosin-containing protein (VCP) (Koller and Brownstein, 1987), an ATPase of mammals and other vertebrates that has an

```
                         Region A                  Region B (ATPase domain)          Region C
                      *    * *** ***          ●  **    ***+*** *******             *  ***
AFG1        50   rddtyqrgIISSLGDLYDSLVKYvpp...yvdVSKIGNSIPRGVYLYGDVGCGKTMLmdlfytti---------PNHLTKKrihf...
VFP-NSF     137  afsvgqqlVFSFNDKLFGLLVKDiea...peiVEQMGCKHVKGILLYGPPGCGKTLLarqigkmlnarepkvvngPEILNKYvges---
SEC18       151  ifsptqylIMEFQGHFFDLKIRNvqa...psvIEKLGISHVKGLLLYGPPGTGKTLIarkigtmlnakepkivngPEILSKYvgss---
CDC48       124  svlpiadtIEGITGNLFDVFLKPyfv...pqlFKAIGIKPPRGVLMYGPPGTGKTLMaravanetgafffling-PEVMSKMages---
VCP-pig     115  hvlpiddtVEGITGNLFEVYLKPyfl...palFKAIGVKPPRGILLYGPPGTGKTLIaravanetgafffling-PEIMSKLages---
PAS1        69   gssenvvlINPVLATVYDLNQKSplv...epiFVNCPLRLRSGILLYGYPGCGKTLLasavaqqcglnfisvkg-PEILNKFigas---
TBP-1       151  sdiggldkqiqelveaivlpmnhkek...---FENLGIQPPKGVLMYGPPGTGKTLLaracaaqtkatflklag-PQLVQMFigdg...

          Region D                    Region E                 Region F
           * *** *               *     *+ +   + *** *  ***    * ** +    *  *  *  *
AFG1     171 EQNLKELGDAKGKEidtvpfl...EIANNSHVLCFDEFQVPDVADAMILR...PPRVAQFTFKQLCGEPLRAEIt----
VFP-NSF  307 EANIRKLFADAEEEqrrlgan...GVEQLNNILVIGMTNRPDLIDEALLR---PGRLEVKMEIGLPDEKGRLQIlhiht
SEC18    320 EENIRNLFKDAEAEyrakgee...DVDQLNNILVIGMTNRKDLIDSALLR---PGRFEVQVEIHLPDEKGRLQIfdiqt
CDC48    292 ESNLRKAFEEAEKNapaiifi...GMNAKKNVFVIGATNRPDQIDPAILR---PGRLDQLIYVPLPDENARLSIlnaql
VCP-pig  283 ESNLRKAFEEAEKNapaiifi...GMSTKKNVFIIGATNRPDIIDPAILR---PGRLDQLIYIPLPDEKSRVAIlkanl
PAS1     776 EQNIRELFERAQSVkpcilff...GAEGLDGVYILAATSRPDLIDSALLR---PGRLDKSVICNIPTESERLDIlqaiv
TBP-1    243 siifidldaigtkrfdsekag...GFQPNTQVKVIAATNRVDILDPALLR---SGRLDRKIEFPMPNEEARARImqihs
```

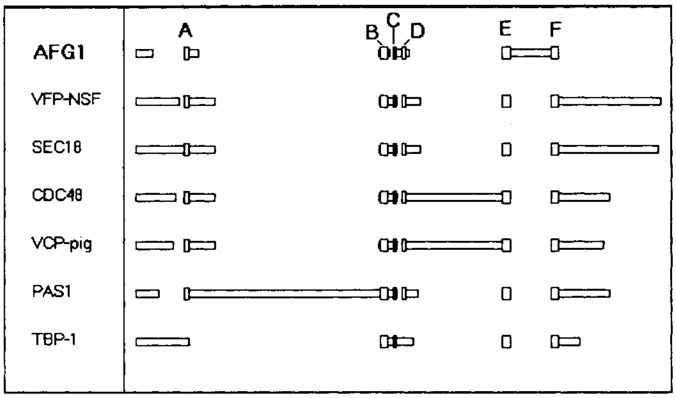

Figure 3. Alignment of AFG1p with proteins in the SEC18–NSF, PAS1, CDC48–VCP and TBP1 group. Region B corresponds to the ATPase domain. Note that TBP1 has no homology with the other proteins in regions A or D. '· · ·' indicates that part of the protein sequence has been omitted, while '– – –' indicates a gap inserted in the sequence for purposes of alignment. (A) Detailed alignment. The most similar residues are indicated by '★' while those that are less so are shown by a '+'. (B) Location of regions aligned in the respective sequences are shown by wide boxes. Note that the aligned regions are only those with homology with AFG1. Narrow boxes are regions that are not aligned with AFG1. Gaps have been introduced to bring the sequences into alignment.

oligomeric structure and is largely soluble (Peters *et al.*, 1990).

The Tat-binding protein 1 (TBP1) is a mammalian nuclear protein whose expression in human cells can suppress the Tat-induced activation of HIV transcription (Neblock *et al.*, 1990).

We describe here a new member of this family, *AFG1* (for ATPase family gene) located on the left arm of chromosome V just centromere-proximal to *MAK10*.

METHODS

Single-stranded templates for DNA sequencing were made from subclones of λ_{6662} in the Stratagene vectors KS^+ and KS^-. Sequencing was by the dideoxy method using Sequenase (US Biochemicals), and the sequence of both strands was determined. The sequence was analysed using the BLAST3 and MACAW programs (Altschul and Lipman, 1990; Schuler *et al.*, 1991).

RESULTS AND DISCUSSION

In sequencing the *MAK10* gene (Lee and Wickner, submitted), we found the beginning of an adjacent open reading frame (ORF) in the 5′ upstream region, which we have now completely sequenced and report here (Figures 1 and 2). The ORF is oriented opposite to *MAK10* in a head-to-head arrangement and

begins 469 nucleotides 5′ to the *MAK10* ORF. The putative protein product is 377 amino acids in length (43 328 Da) with a predicted isoelectric point of 9·92. We name it *AFG1* for reasons described below. Its physical linkage to *MAK10* and the known physical and genetic location of *MAK10* (Lee and Wickner, submitted) show that this gene is just centromere-proximal to *MAK10* on the left arm of chromosome V.

Although a search of the current databases with the FASTA program (Pearson and Lipman, 1988) did not reveal substantial homology to other proteins, the BLAST3 program (Altschul and Lipman, 1990) detected region B of homology of AFG1p with SEC18p, NSF, PAS1p, CDC48p, VCP and TBP1 (Figure 3). This region includes the ATPase motif, G..G.GKT, but there are several other residues in region B that are in common among this group of proteins (Figure 3).

Using MACAW, a program for determining multiple sequence alignments, we found that the similarity between AFG1p and these proteins extends beyond region B to five other regions of the proteins (Figure 3). Searching the databases for proteins with the pattern, 'PD[VIL].D..[VIL]LR' (region E) produced matches only with this group of proteins. The same result was found with the pattern 'P.R[VLF].....L..E..R..I' (region F) or the pattern '[IV]....G.[LFV][YF][DE]..[VIL][KR]' (region A). Since each of these patterns is located in different regions of the AFG1p sequence, this indicates that the similarities shown are significant. While AFG1p is more distantly related to the other proteins in this group than they are to each other, it clearly is homologous. Apart from the putative ATPase domain, the functional significance of this relationship, however, awaits a clearer understanding of the detailed structure–function relationship of these proteins.

REFERENCES

Altschul, S. F. and Lipman, D. J. (1990). Protein database searches for multiple alignments. *Proc. Natl. Acad. Sci. USA* **87,** 5509–5513.

Clary, D. O., Griff, I. C. and Rothman, J. E. (1990). SNAPs, a family of NSF-attachment proteins involved in intracellular membrane fusion in animals and yeast. *Cell* **61,** 709–721.

Eakle, K. A., Bernstein, M. and Emr, S. D. (1988). Characterization of the yeast secretion machinery: identification of the *SEC18* gene product. *Mol. Cell. Biol.* **8,** 4098–4109.

Erdmann, R., Wiebel, F. F., Flessau, A., Rytka, J., Beyer, A., Frohlich, K.-U. and Kunau, W.-H. (1991). *PAS1*, a yeast gene required for peroxisome biogenesis, encodes a member of a novel family of putative ATPases. *Cell* **64,** 499–510.

Frohlich, K.-U., Fries, H.-W., Rudiger, M., Erdmann, R., Botstein, D. and Mecke, D. (1991). Yeast cell cycle protein CDC48p shows full-length homology to the mammalian protein VCP and is a member of a protein family involved in secretion, peroxisome formation, and gene expression. *J. Cell Biol.* **114,** 443–453.

Kaiser, C. A. and Schekman, R. (1990). Distinct sets of *SEC* genes govern transport vesicle formation and fusion early in the secretory pathway. *Cell* **61,** 723–733.

Koller, K. J. and Brownstein, M. J. (1987). Use of a cDNA to identify a supposed precursor protein containing valosin. *Nature* **325,** 542–545.

Moir, D., Steward, S. E., Osmond, B. C. and Botstein, D. (1982). Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties and pseudoreversion studies. *Genetics* **100,** 547–563.

Neblock, P., Dillon, P. J., Perkins, A. and Rosen, C. A. (1990). A cDNA for a protein that interacts with the human immunodeficiency virus Tat transactivator. *Science* **248,** 1650–1653.

Novick, P., Ferro, S. and Schekman, R. (1981). Order of events in the yeast secretory pathway. *Cell* **25,** 461–469.

Pearson, W. R. and Lipman, D. J. (1988). Improved tools for biological sequence analysis. *Proc. Natl. Acad. Sci. USA* **85,** 2444–2448.

Peters, J.-M., Walsh, M. J. and Franke, W. W. (1990). An abundant and ubiquitous homo-oligomeric ring-shaped ATPase particle related to the putative vesicle fusion proteins Sec18p and NSF. *EMBO J.* **9,** 1757–1767.

Riles, L., Dutchik, J. E., Baktha, A., McCauley, B. K., Thayer, E. C., Leckie, M. P., Braden, V. V., Depke, J. E. and Olson, M. V. (1992). Physical maps of the six smallest chromosomes of *Saccharomyces cerevisiae* at a resolution of 2·6 kilobase pairs. In press.

Schuler, G. D., Altschul, S. F. and Lipman, D. J. (1991). A workbench for multiple alignment construction analysis. *Proteins Struct. Funct. Genet.* **9,** 180–190.

Weidmann, P. H., Melancon, P., Block, M. R. and Rothman, J. E. (1989). Binding of an N-ethylmaleimide-sensitive fusion protein to Golgi membranes requires both a soluble protein(s) and an integral membrane receptor. *J. Cell Biol.* **108,** 1589–1596.

Wilson, D. W., Wilcox, C. A., Flynn, G. C., Chen, E., Kuang, W.-J., Henzel, W. J., Block, M. R., Ullrich, A. and Rothman, J. E. (1989). A fusion protein required for vesicle-mediated transport in both mammalian cells and yeast. *Nature* **339,** 355–359.