19. S. Y. M. Lau *et al.*, *J. Biol. Chem.* **259**, 13253 (1984); R. S. Hodges *et al.*, **256**, 1214 (1981).
20. H. Edelhoch, *Biochemistry* **6**, 1948 (1967).
21. L. G. Presta and G. D. Rose, *Science* **240**, 1591 (1988).
22. B. V. Prasad and P. Balaram, *CRC Crit. Rev. Biochem.* **16**, 307 (1989); C. Toniolo and E. Benedetti, *ISI Atlas of Science*, p. 225 (1988).
23. H. R. Wolfe and R. Wilke, *Peptide Res.* **2**, 352 (1989).
24. J. Schneider and S. B. H. Kent, *Cell* **54**, 363 (1988); W. F. DeGrado and J. D. Lear, *Biopolymers* **29**, 205 (1990).
25. Y.-H. Chen, J. T. Yang, K. H. Chau, *Biochemistry* **13**, 3350 (1974); N. Greenfield and G. D. Fasman, *ibid.* **8**, 4108 (1969).
26. R. J. Pollet, B. A. Haase, M. L. Standaert, *J. Biol. Chem.* **254**, 30 (1979).
27. C. N. Pace, *Methods Enzymol.* **131**, 266 (1986); J. U. Bowie and R. T. Sauer, *Biochemistry* **28**, 7139 (1989).
28. These free energy differences refer to the values expected at 300°C, based on an examination of the van't Hoff plots in (*9*).
29. R. W. Williams, A. Chang, D. Juretić, S. Loughran, *Biochem. Biophys. Acta.* **916**, 200 (1987).
30. P. Y. Chou and G. D. Fasman, *J. Mol. Biol.* **74**, 263 (1973); *Biochemistry* **13**, 211 (1974).
31. M. H. Hecht, J. M. Sturtevant, R. T. Sauer, *Proteins* **1**, 43 (1986).
32. S. Dao-Pin, W. A. Baase, B. W. Matthews, *ibid.* **7**, 198 (1990).
33. M. Matsumura, M. Becktel, B. W. Matthews, *J. Biol. Chem.* **264**, 16059 (1989); J. T. Kellis, Jr., K. Nyberg, A. R. Fersht, *Biochemistry* **28**, 4914 (1989).
34. The difference between $\Delta\Delta G_\alpha$ for residues in the center of the helices versus in nonhelical conformations was calculated from the equation $\gamma = \Sigma[f_\alpha(i)(\Delta\Delta G_\alpha)] - \Sigma[f_{n\alpha}(i)(\Delta\Delta G_\alpha)]$, where Σ refers to the sum over all 20 naturally occurring amino acids, and $f_\alpha(i)$ and $f_{n\alpha}(i)$ refer to the probability of finding a given type of amino acid in the center of a helix, or a nonhelical conformation, respectively. The quantities $f_\alpha(i)$ and $f_{n\alpha}(i)$ for all 20 amino acids were calculated from the data compiled in (*4*).
35. J. U. Bowie, J. F. Reidhaar-Olson, W. A. Lim, R. T. Sauer, *Science* **247**, 1306 (1990); R. Kauzmann, *Adv. Prot. Chem.* **14**, 1 (1959).
36. L. Piela, G. N. Némethy, H. A. Scheraga, *Biopolymers* **26**, 1587 (1987).
37. ———, *ibid.*, p. 1273; M. J. McGregor, S. A. Islam, M. J. E. Sternberg, *J. Mol. Biol.* **198**, 295 (1987); J. Janin, S. Wodak, M. Levitt, B. Maigret, *ibid.* **125**, 357 (1978).
38. J. Wojcik, K.-H. Altmann, H. A. Scheraga, *Biopolymers* **30**, 12 (1990).
39. We gratefully acknowledge J. Bryson for his preliminary work on this project. We thank H. A. Scheraga for sharing unpublished data, R. Houghton for peptide synthesis, T. Devine for peptide purification, L. Janvier for amino acid analysis, and J. Lazaar for mass spectrometry.

1 May 1990; accepted 5 September 1990

Protein Splicing Converts the Yeast *TFP1* Gene Product to the 69-kD Subunit of the Vacuolar H^+-Adenosine Triphosphatase

PATRICIA M. KANE, CARL T. YAMASHIRO, DAVID F. WOLCZYK, NORMA NEFF, MARK GOEBL, TOM H. STEVENS*

The *TFP1* gene of the yeast *Saccharomyces cerevisiae* encodes two proteins: the 69-kilodalton (kD) catalytic subunit of the vacuolar proton-translocating adenosine triphosphatase (H^+-ATPase) and a 50-kD protein. The 69-kD subunit is encoded by the 5′ and 3′ thirds of the *TFP1* coding region, whereas the 50-kD protein is encoded by the central third. Evidence is presented that both the 69-kD and 50-kD proteins are obtained from a single translation product that is cleaved to release the 50-kD protein and spliced to form the 69-kD subunit.

VACUOLAR-TYPE PROTON-TRANSLOCATING ADENOSINE TRIphosphatases (H^+-ATPases), which acidify certain intracellular compartments in eukaryotic cells (*1, 2*), have been purified from various sources and show an overall structural similarity (*1*). All of the enzymes purified are multisubunit complexes containing at least two peripheral membrane subunits with molecular masses of ~70 and 60 kD and at least one integral membrane subunit of ~15 kD. The 70-kD subunit contains the catalytic adenosine triphosphate (ATP) binding site. The vacuolar H^+-ATPase of the yeast *Saccharomyces cerevisiae* consists of eight different subunits, including a 69-kD catalytic subunit and 60- and 17-kD subunits similar to those found in other cells (*3, 4*). Genes encoding the 70- and 60-kD subunits have been cloned and characterized from a number of different species, and the sequences of both subunits from plant, animal, and fungal sources are remarkably conserved (*5–7*). The predicted amino acid sequence of the yeast 60-kD subunit is 82 percent identical to that of the *Neurospora crassa* 57-kD subunit and 74 percent identical to that of the human 60-kD subunit (*8*). The functional roles of the yeast vacuolar H^+-ATPase have been assessed by disrupting the gene for the 60-kD subunit, which is termed *VAT2* (*8*). Cells lacking *VAT2* grow more slowly than wild-type cells, fail to grow at neutral pH, and lack the ability to acidify their vacuoles. Isolated vacuoles from these cells lack ATPase activity.

We now present evidence that the 69-kD subunit of the yeast vacuolar H^+-ATPase is one of two proteins encoded by the previously identified *TFP1* gene (*9*). The vacuolar H^+-ATPase subunit is specified by the two ends of *TFP1*, and the central one-third encodes a 50-kD "spacer" protein. Our results indicate that the 69- and 50-kD proteins are formed from a single translation product by post-translational cleavage and splicing.

Encoding of the catalytic subunit of the yeast vacuolar H^+-ATPase by *TFP1*. A dominant allele of *TFP1* was previously isolated from a mutant yeast strain resistant to the drug trifluopera-

P. M. Kane, C. T. Yamashiro, D. F. Wolczyk, and T. H. Stevens are at the Institute of Molecular Biology, University of Oregon, Eugene, OR 97403. P. M. Kane is now in the Department of Chemistry, College of William and Mary, Williamsburg, VA 23185 and C. T. Yamashiro is in the Department of Biological Sciences, Stanford University, Stanford, CA 94305. N. Neff is at the Memorial Sloan-Kettering Cancer Center, New York, NY 10021. M. Goebl was in the Department of Genetics, University of Washington, Seattle, WA 98195, but is now in the Department of Biochemistry, Indiana University School of Medicine, Indianapolis, IN 46202.

*To whom correspondence should be addressed.

zine (*9*). The sequence of wild-type *TFP1* contains an open reading frame that encodes a putative 1071–amino acid, 119-kD protein (*9–11*). Initial analysis of the predicted amino acid sequence revealed homology to the catalytic (β) subunits of H^+-ATPases from *Escherichia coli*, mitochondria, and chloroplasts (F_1F_0-ATPases); further analysis uncovered a much higher degree of identity with the catalytic subunits of vacuolar H^+-ATPases from several different species. As illustrated by the comparison to the *vma-1* gene from *Neurospora* (*5*) (Fig. 1A), the regions of homology mapped to the amino-terminal and carboxyl-terminal thirds (the "N domain" and "C domain," respectively) of the putative protein encoded by the *TFP1* open reading frame. These two regions are 73 and 77 percent identical, respectively, to the corresponding regions of the *Neurospora* 70-kD subunit. The homology to the *Neurospora* protein is interrupted abruptly by the middle one-third of the *TFP1*-encoded protein. This segment (the "spacer domain") shows no amino acid homology to any known vacuolar H^+-ATPase subunit sequence, but is 31 percent identical to one region of the yeast HO endonuclease (*10, 12*). The interruption of homology to the *Neurospora* catalytic subunit occurs in a very highly conserved sequence that is thought to comprise part of the ATP binding site in the catalytic subunits of both vacuolar and F_1F_0 ATPases (*5, 7*).

Additional evidence supports the assignment of *TFP1* as the structural gene for the 69-kD subunit of the yeast vacuolar H^+-ATPase. Amino-terminal protein sequence (17 residues) of the yeast 69-kD subunit was obtained and found to correspond to amino acid residues 13 to 29 of the protein encoded by the *TFP1* open reading frame. This suggests that *TFP1* encodes the 69-kD subunit, but that the first 12 amino acids predicted by the DNA sequence have been removed in the mature protein. [The mature 60-kD subunit of the yeast vacuolar H^+-ATPase also appears to lack the first 12 amino acids predicted by its DNA sequence (*8*).] Analysis of mutant yeast strains lacking *TFP1* provided further evidence that *TFP1* is the structural gene for the 69-kD subunit. The genomic copy of *TFP1* was deleted by replacing the Eco RV–Sph I fragment of *TFP1* (Fig. 1B) with the *LEU2* gene in vitro and then integrating this construction at the *TFP1* locus of haploid yeast cells [*TFP1* is not an essential yeast gene (*9, 10*)] by a single-step gene replacement (*13, 14*). The phenotypes of the resulting *tfp1-Δ8* mutant were compared with the isogenic parental haploid strain. Immunoblots of whole cell lysates indicated that the 69-kD subunit is completely missing from the *tfp1-Δ8* strain, whereas the 60-kD subunit is present in similar amounts in both strains (Fig. 2). Other phenotypes of the mutant cells were identical to those seen for cells lacking *VAT2* (*8*). The *tfp1-Δ8* mutants displayed pH-dependent growth (cell growth was optimal in medium buffered to pH 5 in contrast to no growth in buffered pH 7 medium), and the mutant cells grew more slowly than the isogenic wild-type cells under all conditions. Vacuoles isolated from *tfp1-Δ8* cells had no ATPase activity and were devoid of both the 60- and the 69-kD ATPase subunits, suggesting that cells lacking the 69-kD subunit failed to assemble the 60-kD subunit onto the vacuole membrane. All of the mutant phenotypes could be fully complemented by introducing wild-type *TFP1* on a single-copy (CEN) plasmid. These data, together with evidence presented below, indicate that the 69-kD subunit of the vacuolar H^+-ATPase is encoded by *TFP1* and that this is the sole gene for this subunit.

Comparison of the *TFP1* sequence to that of the *Neurospora vma-1* gene suggested that the mature 69-kD subunit in yeast is composed of the N and C domains of the *TFP1* gene product, and that the spacer domain is removed at some point in the synthesis of the subunit. Support for this structural model was obtained by mapping the epitopes recognized by four different monoclonal antibodies that reacted with the 69-kD subunit on immunoblots (*15*). Fragments of *TFP1* were placed under an inducible promoter and expressed in *Escherichia coli*. We then assessed the ability of the different antibodies to recognize the proteins expressed on induction by comparing lysates from induced and uninduced cells on immunoblots (*16*). From this analysis, we determined that three of the monoclonal antibodies recognized epitopes in the N domain of the *TFP1* gene product and one recognized an epitope in the C domain. None of the four antibodies recognized the spacer domain (as defined by the Kpn I–Nae I fragment of *TFP1*). These results are consistent with the idea that the N and C domains are joined to form the 69-kD subunit.

***TFP1* spacer domain region present in the mature mRNA.** RNA splicing represents one obvious way of joining the N and C domains of the *TFP1* gene product. We addressed this possibility by RNA (Northern) analysis of the *TFP1* mRNA (Fig. 3). Polyadenylated RNA [poly(A)$^+$ RNA] was isolated and probed with DNA fragments corresponding exclusively to the N, spacer, and C domains. All of these probes recognized a single mRNA of 3.7 kilobases (kb), which is large enough to encode the three domains of the predicted polypeptide. No specific transcripts of the expected size for an mRNA processed to remove the region encoding the spacer domain were seen. Although the size of the message and the recognition of the same species by all three probes suggested that it contains the entire *TFP1* sequence, we also analyzed the size of the mRNA produced from an allele of *TFP1* from which the region corresponding to the spacer domain had been precisely deleted by oligonucleotide-directed mutagenesis. Cells carrying this *tfp1-Δ10* allele synthesize a functional 69-kD subunit (described below) and should produce a transcript comparable in size to a spliced mRNA encoding only the N and C domains. As expected, this construction produced an mRNA of ~2.4 kb (*17*). This experiment does not eliminate the possibility that an undetectably small amount of spliced message is present and is responsible for synthesis of the 69-kD subunit. However, other evidence argues against splicing of the *TFP1* mRNA. Introns in yeast nuclear genes contain highly conserved sequences at the 5′ end and at the branch site (*18*), and neither of these sequences is present in *TFP1*. Indeed, *rna2*^{ts} *rna8*^{ts} double mutants, which are defective in nuclear mRNA splicing at the nonpermissive temperature (*19*), produce the normal 69-kD subunit (and no aberrant products from unspliced message) even after long times at the restrictive temperature (*20*). Group I and group II introns also contain conserved sequences that form second-

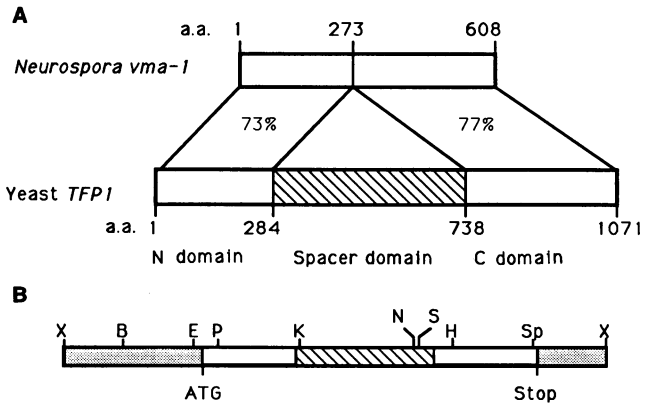

Fig. 1. (**A**) Schematic comparison of yeast *TFP1* and *Neurospora vma-1* genes. Amino acid (a.a.) sequences were compared and percentage identities between predicted amino acids 1 to 608 from the *Neurospora* gene and 20 to 284 and 738 to 1071 from yeast *TFP1* are shown. (**B**) Restriction map of *TFP1*. The open and slashed boxes correspond to the same regions as in (A); the dotted boxes are the 5′ and 3′ untranslated regions. Restriction enzymes: X, Xba I; B, Bam HI; E, Eco RV; H, Hind III; K, Kpn I; N, Nae I; P, Pst I; S, Sac II; Sp, Sph I. The overall restriction map of the gene matched that reported previously for the *TFP1* dominant allele exactly (*9*).

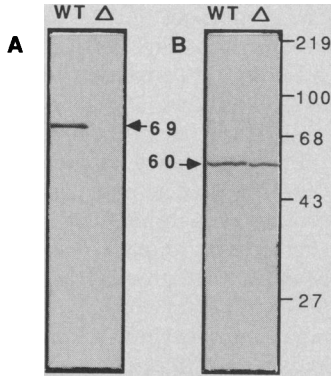

Fig. 2. Immunoblots of whole cell lysates from wild-type and *tfp1-Δ8* cells. Yeast were grown to mid-logarithmic phase in yeast extract peptone dextrose (YEPD) medium buffered to pH 5.0 with 50 mM sodium phosphate and 50 mM sodium succinate. Total cell protein extracts were prepared, separated by SDS–polyacrylamide gel electrophoresis (PAGE) on a 10 percent acrylamide gel, and blotted onto nitrocellulose (*8*). Extracts corresponding to an equal number of cell equivalents of SF838-5A (WT) or *tfp1-Δ8* (Δ) cells were probed with a mixture of monoclonal antibodies 11E6 and 8B1 to the 69-kD subunit (**A**) or monoclonal antibody 13D11 to the 60-kD subunit (**B**) of the yeast vacuolar H^+-ATPase (*15*). The positions of the molecular mass standards are shown in kilodaltons.

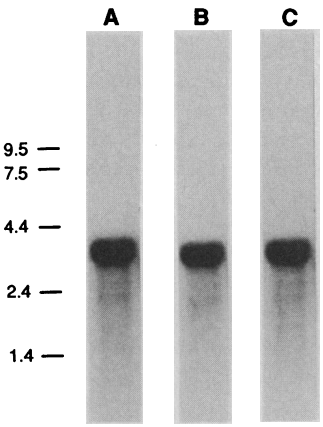

Fig. 3. RNA (Northern) blot analysis of *TFP1* transcripts. Polyadenylated [poly(A)$^+$] mRNA was isolated from wild-type yeast cells (*30*), fractionated, and transferred to a nylon filter (*32*). Northern blots were probed with ^{32}P-labeled fragments of *TFP1* corresponding to (**A**) the N domain (Eco RV–Pst I fragment), (**B**) spacer domain (Kpn I–Sac II fragment) and (**C**) the C domain (Hind III–Xba I fragment) (*9*). Positions of RNA molecular size markers (BRL) are shown in kilobases.

ary structures important for their splicing (*21*), and the region encoding the spacer domain shows neither the conserved sequences nor any obvious secondary structure. Finally, the presence of a single, long open reading frame is very unusual for intron-containing genes.

The presence of the region encoding the spacer domain in the mRNA led us to look for a large protein precursor of the 69-kD subunit. Overexpression of *TFP1* might saturate a step in the processing of the *TFP1* gene product, permitting visualization of a precursor. *TFP1* was placed under control of the inducible yeast *GAL10* promoter (*22, 23*) (Fig. 4). Lysates from cells grown 6 hours in the presence or absence of galactose were probed with antibodies specific for the N domain or C domain of the 69-kD subunit. The lysates from the cells grown in galactose show overproduction of the 69-kD subunit, but they also show a very small amount of a 119-kD product that is recognized by both antibodies. However, although this protein is the size expected for an unprocessed precursor containing all three domains encoded by *TFP1*, we were never able, because of its low abundance, to prove from kinetic (pulse-chase) experiments that it converted to the 69-kD subunit (see below).

The spacer domain of the *TFP1* gene product must be translated to obtain the mature 69-kD subunit. The foregoing experiments indicated that the region corresponding to the spacer domain is present in the mature mRNA. However, these experiments did not directly address the question of whether this region must be translated in order to generate the 69-kD subunit. We addressed this question by making a series of mutations in the spacer domain region of *TFP1* and determining whether CEN plasmids containing these mutant alleles could complement the defects of a *tfp1-Δ8* strain (*23*). Deletion of four nucleotides at the Kpn I site (*tfp1-Δ14*) or two nucleotides at the Sac II site (*tfp1-Δ12*) of the spacer domain region introduces a stop codon within 18 codons of the respective restriction sites (Fig. 5A). The *tfp1-Δ8* cells containing a plasmid-borne copy of *TFP1* produce a normal 69-kD subunit, as expected (Fig. 5B). In contrast, no protein cross-reactive with the monoclonal antibodies to the 69-kD subunit that recognize the N domain was detected in cells containing the *tfp1-Δ14* allele. Cells carrying the *tfp1-Δ12* plasmid contained only a small amount of cross-reactive material, which migrated at an apparent molecular mass of 72-kD, suggesting formation of a truncated protein composed of the N domain of the 69-kD subunit and a part of the spacer-domain terminating after the Sac II site. Consistent with this, antibodies recognizing only the C domain did not cross-react with the 72-kD product (*20*). These results indicate that continuous translation of the entire *TFP1* mRNA is required for synthesis of the 69-kD H^+-ATPase subunit.

We took this proposal one step further. If these mutant *TFP1* alleles fail to produce the 69-kD subunit because of stop codons introduced in the spacer region, then introduction of a second mutation that restores the reading frame before the stop codon is reached might also restore production of the 69-kD subunit. We tested this hypothesis by introducing an additional 1-bp (base pair) deletion three nucleotides downstream of the original 2-bp deletion in *tfp1-Δ12* mutation by oligonucleotide-directed mutagenesis to produce a 3-bp deletion (*tfp1-Δ13*) (Fig. 5A). A plasmid containing the *tfp1-Δ13* allele restored normal growth at pH 7 to *tfp1-Δ8* cells and abolished the other phenotypes of the *tfp1-Δ8* mutants. Cells carrying the *tfp1-Δ13* allele produced an apparently normal 69-kD subunit (Fig. 5B). This result suggests that the mutant phenotypes seen with the original *tfp1-Δ12* mutation are due to the shift in reading frame and cannot be attributed to disruption of sites important for RNA processing. It also provides further support for mechanisms requiring continuous translation of the entire *TFP1* open reading frame.

To determine whether the presence of the spacer domain region was essential for synthesis of the 69-kD subunit, we deleted the entire spacer region, fusing the N and C domain regions of *TFP1* (*tfp1-Δ10*). A plasmid with this construction restored normal growth at pH 7 to *tfp1-Δ8* cells and directed the synthesis of a 69-kD subunit indistinguishable from the wild-type subunit (Fig. 5B). This result indicates that the N and C domains can form a functional subunit when joined by a normal peptide bond. Because this fusion of the N and C domains creates a functional subunit that has an apparent

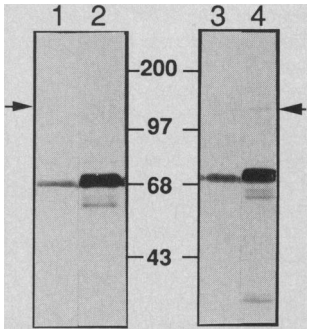

Fig. 4. Immunoblots of yeast cells carrying *TFP1* under control of an inducible *GAL10* promoter. CJRY20-3B cells were transformed with the plasmid pPK10, which contains *TFP1* fused to the yeast GAL10 promoter (*22*). Cells were grown overnight in medium containing 2 percent raffinose as described (*37*) and induced by the addition of 2 percent galactose directly to the growth medium. Uninduced cells had no galactose added. Cells were harvested 6 hours after induction, and total protein extracts were prepared, separated by SDS-PAGE on a 10 percent acrylamide gel, and blotted onto nitrocellulose (*8*). Extracts corresponding to an equal number of cell equivalents of induced (lanes 2 and 4) or uninduced (lanes 1 and 3) cells were probed either with monoclonal antibody 8B1, which recognizes the N domain (lanes 1 and 2), or with monoclonal antibody 11E6, which recognized the C domain (lanes 3 and 4). Arrows indicate the position of the 119-kD product that represents the unprocessed *TFP1* gene product in the induced cells. The positions of molecular mass standards are shown in kilodaltons.

molecular mass identical to the wild-type protein, our proposed primary structure for the 69-kD subunit, based on homology to the *Neurospora vma-1* gene, must be reasonably close to the primary structure of the actual protein.

These observations prompted us to investigate whether the spacer domain region of *TFP1* encoded a stable protein. We injected rabbits with a peptide from the predicted amino acid sequence of the spacer domain (*15*), and then used the resulting antibodies to probe immunoblots of wild-type cells and *tfp1-Δ8* cells carrying various plasmids. The specific antibodies to the peptide recognized a protein of ~50-kD in both wild-type cells and *tfp1-Δ8* cells carrying a CEN-*TFP1* plasmid (Fig. 5C). This protein was missing in *tfp1-Δ8* cells and in the *tfp1-Δ10* cells. The antibodies recognized the same protein of ~72-kD that cross-reacted with the antibodies to the 69-kD subunit (N domain) in the *tfp1-Δ12* cells, consistent with the formation of a truncated protein from this construction. These results indicate that the spacer domain region of *TFP1* encodes a fairly abundant and stable protein. However, the immunoblots may not accurately reflect the relative proportions of the 69-kD (ATPase subunit) and 50-kD (spacer) products and would not reveal transient precursors in the formation of these products.

The biosynthesis of the two protein products from *TFP1* was then examined. Cells were labeled for 10 minutes with [^{35}S]methionine and [^{35}S]cysteine, and either lysed immediately or incubated for 10 minutes more in the presence of unlabeled methionine and cysteine before lysis (Fig. 6). The newly synthesized proteins were immunoprecipitated with either the antibodies to the 69-kD subunit or the antibodies to the spacer peptide. The similar amounts of protein precipitated by the two different sets of antibodies indicate that the 69- and 50-kD proteins are produced at similar rates. Both products must be fairly stable because there was no obvious loss of product after the 10-minute incubation period. No full-length (119-kD) translation product was detected in this experiment or with labeling times as short as 2 minutes. Similar time-labeling studies on yeast cells that had been induced to overexpress *TFP1* revealed a minor amount of a 119-kD product relative to the 69-kD product (*20*).

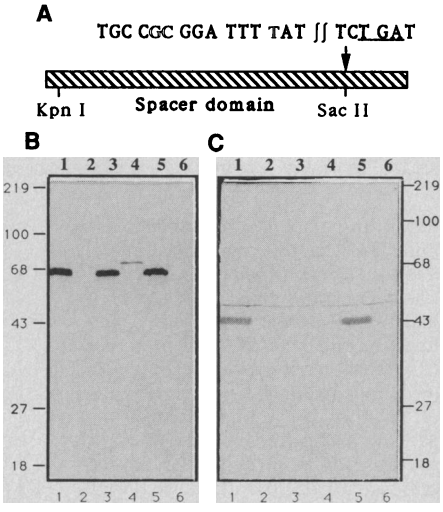

Fig. 5. (**A**) Mutations at the Sac II site. Highlighted nucleotides correspond to nucleotides removed in the two different mutations, the break (∫∫) covers 48 nucleotides, and the stop codon brought into frame by the 2-bp deletion is underlined. (**B** and **C**) Immunoblots of cells containing *TFP1* spacer domain mutations on CEN plasmids. *tfp1-Δ8* cells were transformed by the lithium acetate method (*29*). Transformed cells were selected and maintained on supplemented SD medium lacking uracil. Whole cell lysates were prepared, subjected to SDS-PAGE, and blotted as described (Fig. 2). Blots were probed with a mixture of monoclonal antibodies 8B1 and 11E6 to the 69-kD subunit (**B**) and with the antibodies to the spacer peptide (*15*) (**C**). An identical blot probed with the monoclonal antibody to the 60-kD subunit showed an equal amount of protein of the expected molecular size in all the samples (20). Cells were transformed with the following plasmids: (lane 1) pDW21 (CEN-*TFP1*); (lane 2) pRS316 (vector only); (lane 3) ptfp1-Δ10 (spacer Δ); (lane 4) ptfp1-Δ12 (Sac II Δ2); (lane 5) ptfp1-Δ13 (Sac II Δ2 + Δ1); (lane 6) ptfp1-Δ14 (Kpn I Δ4). The positions of molecular mass standards are shown in kilodaltons.

The proportions of the two products were similar to those seen at steady-state in Fig. 4, even at short labeling times, and it thus was not possible to establish whether the 119-kD "precursor" converted to the 69-kD and 50-kD proteins with time because it represented a small percentage of the immunoprecipitated protein.

Protein splicing of the *TFP1* gene product in heterologous contexts. To determine whether production of the 69- and 50-kD proteins from the *TFP1* mRNA requires factors specific to yeast, we attempted to duplicate this process in vitro. *TFP1* was cloned behind an SP6 promoter, and RNA was synthesized in vitro by the SP6 RNA polymerase. The synthesized RNA was then added to a rabbit reticulocyte lysate translation system containing [^{35}S]methionine and [^{35}S]cysteine, and the products were immunoprecipitated with either the antibodies to the 69-kD subunit or the antibodies to the spacer peptide. The antibodies to the 69-kD subunit immunoprecipitated a 69-kD product and antibodies to the spacer peptide precipitated a similar amount of 50-kD product (Fig. 7). The appearance of these two products was absolutely dependent on addition of the *TFP1* RNA because nothing was precipitated by either antibody in the absence of RNA. As in the immunoprecipitations of *TFP1* products synthesized in vivo (Fig. 6), no full-length 119-kD translation product was seen. Similar results were obtained with mRNA isolated from yeast (*20*). These results indicate that (i) there is nothing specific to the yeast mRNA that is required for production of the 69- and 50-kD proteins and (ii) either the cellular machinery for synthesis and splicing of the 69- and 50-kD proteins is preserved in mammals and present in the reticulocyte translation mixture or the process of forming the 69- and 50-kD products requires no cellular machinery and is therefore autocatalytic.

We used *TFP1* expression in *E. coli* to determine how the gene product would be handled in bacteria. *TFP1* was cloned into the

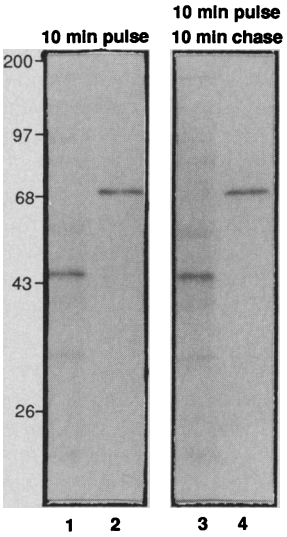

Fig. 6. Synthesis of the 69-kD ATPase subunit and the 50-kD spacer protein. SF838-5A cells were grown to mid-logarithmic phase in supplemented SD medium lacking methionine. Cells were converted to spheroplasts with zymolyase 100T (ICN Biochemicals) as described (*38*). Spheroplasts were resuspended at a density of 3.3×10^7 per milliliter in growth medium containing 1.2 M sorbitol and incubated with 200 μCi of Tran^{35}S-label (ICN Biochemical) per 10^7 cells at 30°C for 10 minutes (pulse). At the end of the labeling period, half of the culture was combined with an equal volume of ice-cold 20 percent trichloroacetic acid (TCA) (lanes 1 and 2). Unlabeled methionine and cysteine were added to the other half of the culture to a final concentration of 0.3 mg/ml and the incubation was continued for an additional 10 minutes (lanes 3 and 4), followed by TCA precipitation (chase). The TCA-precipitate was washed once in acetone and solubilized in 0.9 M tris-HCl (pH 8.0), 1 percent SDS, 1 percent Triton X-100, and 20 mM EDTA (immunoprecipitation buffer) at 100°C. The solubilized pellets were diluted (1:10), then immunoprecipitated with either a mixture of monoclonal antibodies 8B1 and 11E6 (lanes 2 and 4) or the antibodies to the spacer peptide (lanes 1 and 3). For the spacer peptide antibody samples, the initial incubation with primary antibody was followed by incubations with goat antibody to rabbit immunoglobulin G (IgG) and then mouse antibody to goat IgG (both from Jackson Immunobiologicals) to increase the stability of the immune complexes. For all of the immunoprecipitations, the immune complexes were precipitated with goat antibody to mouse IgG conjugated to Sepharose-4B (Organon-Teknika). The precipitated samples were washed twice in immunoprecipitation buffer without SDS, then analyzed by SDS-PAGE and fluorography as described (*39*). The positions of molecular mass standards are shown in kilodaltons.

pCW-ori+ vector behind the *tac* promoter. Three identical immunoblots (Fig. 8) of *E. coli* cell lysates made 15 minutes after isopropylthiol-β-D-galactoside (IPTG) induction were probed with antibodies to the N, spacer, and C domains. Once again, there was no 119-kD product visible, but the antibodies to the spacer peptide recognize a protein of 50 kD and the antibodies to the 69-kD subunit recognized a 69-kD protein. This implies not only that the gene product can be correctly cleaved in *E. coli*, but also that splicing is taking place, because the 69-kD product is recognized by antibodies to the N and C domains but not by antibodies to the spacer domain. None of the antibodies recognized any protein in immunoblots of lysates from uninduced cells (*20*). At later times of induction (3 hours), a small amount of 119-kD product that was recognized by all three sets of antibodies was formed. This probably represents the complete *TFP1* translation product, but was consistently a minor species relative to the 69- and 50-kD products. The antibody to the C domain also recognized a stable protein of 38 kD. It is unclear whether this is a stable fragment resulting from proteolysis of *TFP1* product, a genuine intermediate in the processing-splicing reaction, or a product of incorrect initiation of translation within the *TFP1* mRNA. However, it is intriguing that this fragment is approximately the predicted size of the C domain (37.1 kD) and is recognized only by the antibody to the C domain. The production of the 69- and 50-kD proteins in *E. coli* is more consistent with an autocatalytic model for protein splicing of the *TFP1* gene product.

The 69-kD and 50-kD proteins from a single translation product by protein splicing. Our data thus invite two interpretations. The 69- and 50-kD proteins could be produced by two different mRNA molecules that either are present in very small amounts or are very short-lived (and therefore undetectable on the Northern blots). Alternatively, the two proteins could be produced from a single translation product that is cleaved to release the spacer protein and then "spliced" to form the mature 69-kD subunit. We believe that all the data taken together favor the second possibility. The immunoprecipitation data indicate that the two proteins are produced at nearly identical rates. Although it is possible that two different mRNA's could be produced and then translated at the same rate, this would seem to be less likely than synthesis of a single product and division into two parts. As discussed above, the splicing of the mRNA would also have to involve previously uncharacterized mechanisms because none of the consensus sequences for RNA splicing are present in *TFP1*. Results from the spacer domain mutations indicate that the translation of the 69- and 50-kD products are very closely linked. Introduction of stop codons in the spacer domain region of *TFP1* by the creation of small deletions alters both the 69- and 50-kD products, restoration of the reading frame in the spacer domain region restores both products, and, in the case of the *tfp1-Δ12* allele, translation must continue through the junction of the N and spacer domains in order to generate the truncated product seen. Because conditions were never found that resulted in accumulation of the 119-kD *TFP1* translation product, and because the 69- and 50-kD species predominate even at very short labeling times, it seems likely that the cleavage and splicing steps occur very rapidly. If the processing reactions are autocatalytic as the results shown in Fig. 8 suggest, overproduction of the protein would not be predicted to result in accumulation of a precursor, and the cleavage and splicing reactions might occur before or immediately after release of the newly synthesized protein from the ribosome.

We can envision many variations on these two basic models, but most of these variations are less consistent with our data than the protein splicing model. Hirata *et al.* (*10*) note that a 10–amino acid stretch of the spacer domain of the *TFP1* gene product is also found in many RNA maturases. Although this might suggest a mechanism in which an initial translation product that includes the maturase-like domain produces a processed transcript that encodes the 69-kD subunit (*24*), several lines of evidence are not consistent with such a model. Other genes that use RNA maturases in their processing, such as the cytochrome b gene of *S. cerevisiae* (*25*), show mRNA species corresponding to both the maturase-containing protein product and the processed product. We would also predict that an RNA maturase would act in trans on *TFP1* transcripts, so that *TFP1* alleles containing spacer mutations could produce a normal 69-kD subunit when expressed in cells containing a wild-type *TFP1* gene capable of providing the maturase function. In fact, the results indicate that the processing event does not occur in trans. When the *tfp1-Δ12* allele was expressed in wild-type cells, the cells produced the 72-kD protein product (N domain plus a truncated spacer domain) along with a normal 69-kD subunit, indicating that the wild-type gene cannot compensate for the mutation in the spacer region (*20*). There has been one report of "ribosome skipping," in which a 60-bp section of the mRNA of the p60 protein of T4 topoisomerase was not translated (*26*). However, the skipping appears to be dependent on formation of a hairpin between regions within the segment that is skipped and the 5′ junction, and no similar sequences for secondary structure are apparent at the proposed junctions in *TFP1*. In addition, the translation of *TFP1* produces both the 69- and 50-kD proteins.

We have no direct evidence that the N and C domains are spliced by the formation of a normal peptide bond, but substantial indirect evidence favors this possibility. When the spacer domain was deleted genetically, a functional 69-kD product was formed, and there is no reason to propose that this construction would produce anything other than one continuous polypeptide. If the 69-kD protein produced from the wild-type *TFP1* gene contains another type of bond, then it must be very stable, because the 69-kD subunit is stable to reducing agents, many detergents (including SDS), chaotropic anions, 10 percent trichloroacetic acid, 0.1 M NaOH, and 100°C. When these facts are assessed together, the model based on protein splicing (Fig. 9) is most consistent with all the data.

How exceptional is protein splicing? Production of the 69- and 50-kD proteins from a single translation product by protein splicing would represent a major departure from the usual colinearity of mature mRNA and protein sequences. Nevertheless, the idea of protein splicing is not unprecedented. In concanavalin A, a jack bean lectin, 15 amino acids in the middle of the predicted amino acid

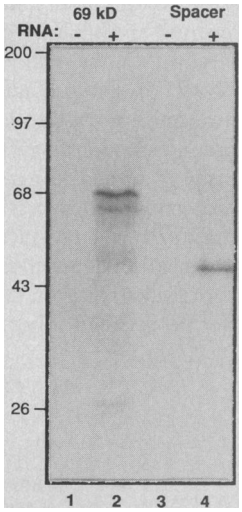

Fig. 7. In vitro transcription and translation of *TFP1*. Plasmid pCY59 was linearized by digestion with Sal I, then transcribed with SP6 RNA polymerase (Promega). The in vitro transcribed RNA was added to a rabbit reticulocyte translation mixture (BRL), and translation was carried out in the presence of 20 μCi of Tran^{35}S-label per reaction. The newly synthesized proteins were immunoprecipitated with the monoclonal antibodies 8B1 and 11E6 to the 69-kD subunit, followed by goat antibodies to mouse IgG conjugated to Sepharose-4B (lanes 1 and 2) as described in Fig. 6 or the antibodies to the spacer peptide followed by IgGSorb (The Enzyme Center) (lanes 3 and 4) as described (*8*). No RNA was added to the translation mixture in the samples shown in lanes 1 and 3. The smaller molecular mass products seen in the 69-kD subunit precipitation appear to lack part of the C domain, because they were recognized by antibodies to the N domain but not the C domain in control experiments. It is unclear if they result from proteolysis or premature termination of transcription or translation. The positions of the molecular mass standards are shown in kilodaltons.

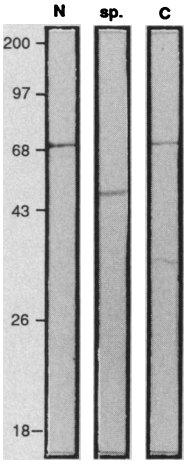

Fig. 8. Splicing of *TFP1* product in *E. coli*. Expression of *TFP1* was induced by the addition of 1 mM IPTG to cells in early logarithmic growth. Cells were harvested 15 minutes after induction and lysed by freezing at −80°C and then boiling in 5 percent SDS and 8 M urea. Blots were prepared as in Fig. 2, then probed with a mixture of monoclonal antibodies to the N domain of the 69-kD subunit (8B1, 7D5, and 7H12) (N), antibodies to the spacer peptide (Sp.), or monoclonal antibody 11E6 to the C domain of the 69-kD subunit (C). The positions of molecular mass standards are shown in kilodaltons.

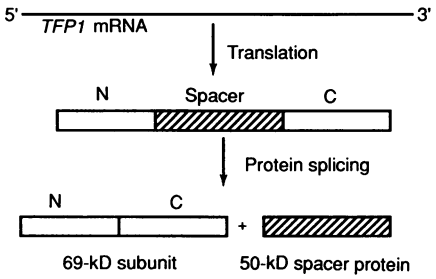

Fig. 9. Model representing protein splicing of the *TFP1* gene product to form the 69-kD and 50-kD products.

sequence (spacer domain) are missing from the mature protein (*27, 28*). Timed labeling studies have revealed that the primary translation product is cleaved and a new peptide bond formed to rejoin the N and C domain fragments into the mature protein. In contrast to the yeast 69-kD subunit, protein splicing of concanavalin A results in the N domain of the primary translation product becoming the C domain of the mature protein. Thus, the N and C domains swap relative positions in the primary translation product and the mature polypeptide. The kinetics of splicing and computer modeling of the mature concanavalin A structure indicate that the splicing occurs by transpeptidation (*27*). In addition to these likely examples of protein splicing, there are many examples where the possibility of protein splicing has simply not been addressed, either because all of the information about the protein is extrapolated from its cDNA sequence or because aberrations in the behavior of the protein, such as unusual mobilities on SDS-polyacrylamide gels, were attributed to trivial causes without further investigation.

Our results provide a framework for the elucidation of the mechanisms involved in protein splicing and for an assessment of its prevalence. Concanavalin A is not found in an organism that can be easily manipulated genetically, so studies on the mechanisms of its processing were necessarily limited. Genetic analysis of processing of the *TFP1* gene product should provide information on the portions of the molecule necessary for its splicing in yeast and may also permit identification of other cellular components that participate in this process and other proteins that are spliced.

REFERENCES AND NOTES

1. M. Forgac, *Physiol. Rev.* **69**, 765 (1989).
2. I. Mellman, R. Fuchs, A. Helenius, *Annu. Rev. Biochem.* **55**, 663 (1986).
3. P. M. Kane, C. T. Yamashiro, T. H. Stevens, *J. Biol. Chem.* **264**, 19236 (1989).
4. E. Uchida, Y. Ohsumi, Y. Anraku, *ibid.* **260**, 1090 (1985).
5. E. J. Bowman, K. Tenney, B. J. Bowman, *ibid.* **263**, 13994 (1988).
6. B. J. Bowman, R. Allen, M. A. Wechser, E. J. Bowman, *ibid.*, p. 14002, M. F. Manolson, B. F. F. Oullette, M. Filion, R. J. Poole, *ibid.*, p. 17987; H. Nelson, S. Mandiyan, N. Nelson, *ibid.* **264**, 1775 (1989); T. C. Südhof, V. A. Fried, D. K. Stone, P. A. Johnston, X.-S. Xie, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6067 (1989).
7. L. Zimniak, P. Dittrich, J. P. Gogarten, H. Kibak, L. Taiz, *J. Biol. Chem.* **263**, 9102 (1988).
8. C. T. Yamashiro, P. M. Kane, D. F. Wolczyk, R. A. Preston, T. H. Stevens, *Mol. Cell. Biol.* **10**, 3737 (1990).
9. C-K. Shih, R. Wagner, S. Feinstein, C. Kanik-Ennulat, N. Neff, *ibid.* **8**, 3094 (1988).
10. R. Hirata *et al.*, *J. Biol. Chem.* **265**, 6726 (1990).
11. During the original sequence determination of *TFP1* (*9*) a single T was missed at nucleotide 961. The addition of this base to the sequence shifts the reading frame of the NH_2-terminal region, resulting in the addition of 40 amino acids to the beginning of the protein.
12. D. W. Russell *et al.*, *Mol. Cell. Biol.* **6**, 4281 (1986).
13. R. K. Rothstein, *Methods Enzymol.* **101**, 202 (1983).
14. Yeast strains used were SF838-5A (*leu2-3,-112*; *ura3-52*; *ade6*; *MAT*α), SF838-5A *tfp1-Δ8* (SF838-5A with *tfp1-Δ8*::*LEU2*), and CJRY20-3B (*ade2-101*, *his3-Δ300*, *ura3-52*, *lys2-801*, *GAL2*, *MAT*α). The *tfp1-Δ8* strain was constructed by transforming SF838-5A cells with the 4.3-kb Xba I fragment of pPK8 by the lithium acetate procedure (*29*). Cells were grown in YEPD medium or synthetic dextrose (SD) medium with the appropriate supplements (*30*).
15. Preparation of monoclonal antibodies to the 69-kD subunit has been described (*3*). The antibody to the 60-kD subunit (13D11) was derived from a separate fusion in which frozen splenocytes from the same mouse were used. Hybridoma supernatant fractions were screened for recognition of subunits of the purified vacuolar H^+-ATPase on protein immunoblots. The antibodies to the spacer peptide were made by immunizing and boosting rabbits with the synthetic peptide: NH_2-RGRET-MYSVVQKSQHRAHK(C)-COOH, corresponding to amino acids 325 to 347 of the *TFP1* gene product. The peptide was conjugated to keyhole limpet hemocyanin with the use of *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester, and a rabbit was injected as described [N. Green *et al.*, *Cell* **28**, 477 (1982)]. Crude serum was affinity purified against the protein product encoded by the Kpn I–Nae I fragment of *TFP1* produced in *E. coli* as described (*31*) and conjugated to cyanogen-bromide activated Sepharose.
16. The epitopes recognized by the monoclonal antibodies to the 69-kD subunit were mapped as follows: the 1.2-kb Kpn I–Nae I fragment of *TFP1* was cloned into the Kpn I and Xba I (blunted) sites of pEXP1 (*31*), the 1.9-kb Nae I-Xba I fragment was cloned into Sma I and Xba I-digested pEXP2 (*31*), and the 1.8-kb Pst I (blunted)–Nae I fragment was ligated into Sma I-digested pEXP2. *Escherichia coli* lysates were prepared before and 3 hours after induction with 2 mM IPTG, and lysates were separated by SDS–polyacrylamide gel electrophoresis (SDS-PAGE) and blotted as described (*8*). Monoclonal antibody 11E6 recognized only the expressed Nae I–Xba I fragment (C domain) and antibodies 8B1, 7D5, and 7H12 recognized nothing in the Kpn I–Nae I fragment (spacer domain), but recognized a band of the expected size in the Pst I–Nae I construction (N plus spacer domains).
17. N. Neff, unpublished data.
18. J. L. Teem *et al.*, *Nucleic Acids Res.* **12**, 8295 (1984).
19. A. J. Lustig, R-J. Lin, J. Abelson, in *Molecular Biology of RNA: New Perspectives*, M. Inouye and B. S. Dudock, Eds. (Academic Press, New York, 1987).
20. P. M. Kane, C. T. Yamashiro, T. H. Stevens, unpublished data.
21. T. R. Cech and B. L. Bass, *Annu. Rev. Biochem.* **55**, 599 (1986); P. A. Sharp, *Science* **235**, 766 (1987).
22. M. Johnston and R. W. Davis, *Mol. Cell. Biol.* **47**, 1440 (1984).
23. Plasmids were constructed by standard DNA manipulations (*32*). pDW21 (CEN-*TFP1*): The 5.5-kb Xba I fragment of *TFP1* was cloned into the Xba I site of the yeast shuttle vector pSEYC68 (*33*). For pTFP1, the 5.5-kb Xba I fragment of *TFP1* was ligated into an Xba I–cut pUC12 vector; for pPK8, pTFP1 was cut with Eco RV and Sph I, and Sph I site was blunted. The Eco RV–Sph I segment of *TFP1* was replaced with the 2.1-kb Hpa I fragment of *LEU2*. The *LEU2* gene was oriented in the opposite direction from the remainder of *TFP1* in the resulting plasmid. The pFUS plasmid is a 2-μm yeast multicopy plasmid containing a *LEU2* selectable marker and the promoter region of *GAL10* upstream from a multiple cloning site (*17*). We prepared pPK10 by inserting the Eco RV–Sal I fragment of pTFP1 into the Nco I and Xho I sites of pFUS after the Nco I site was blunted. The final construction places the *GAL10* promoter at position −91 of *TFP1*. For ptfp1-Δ14 (CEN–*Kpn 1Δ4*), pTFP1 was digested with Kpn I, the 5′ overhang was cut back, and the plasmid was religated. The resulting plasmid was digested with Bam HI and Sal I (Sal I cuts in the polylinker of pUC12, on the 3′ side of *TFP1*) and the 5-kb fragment was ligated to Bam HI and Sal I–digested pSEYC68; for ptfp1-Δ12 (CEN–*Sac IIΔ2*): pDW21 was digested with Sac II, the 5′ overhang was cut back, and the plasmid was religated; for ptfp1-Δ13 (CEN–*Sac IIΔ2* + *Δ1*): The 5-kb Bam HI–Sal I fragment of ptfp1-Δ12 was transferred to the shuttle vector pRS316 (*34*), which has an M13 origin, permitting retrieval of single-stranded DNA. Single-stranded DNA was recovered from a *dut*$^-$*ung*$^-$ *E. coli* strain (CJ236), and in vitro synthesis of the complementary strand was primed by addition of the oligonucleotide: 5′-ACGTGAGTGCCGGATTTATTTCGAGTTACA (*35*); for ptfp-Δ10 (spacerΔ): The 1.2-kb Kpn I–Nae I fragment of pTFP1 was deleted and the 4 kb Bam HI–Sal I fragment of the resulting plasmid was transferred to Bluescript KS^+(Stratagene). Single-stranded DNA was recovered and a complementary mutant strand was synthesized in vitro as described above; the oligonucleotide 5′-CCATTATCTATGTCGGGTGCGGAGAAAGAGGTAATGAAAT-3′ was used as the primer. The 3.8-kb Bam HI–Sal I fragment of the mutant plasmid was ligated to Bam HI–Sal I digested pSEYC68; pCY59: pTFP1 was digested with Eco RV and Sal I, then ligated with pGEM2 that had been cut with Sma I and Sal I. Sequences of the mutant plasmids were confirmed by DNA sequencing by the dideoxy chain termination method (*36*). *E. coli* strains MC1066, JM101, and CJ236 were used for all the DNA manipulations (*32*).
24. P. S. Perlman and R. A. Butow, *Science* **246**, 1106 (1989).
25. J. Lazowska *et al.*, *J. Mol. Biol.* **205**, 275 (1989).

26. W. M. Huang *et al.*, *Science* **239**, 1005 (1988).
27. D. J. Bowles *et al.*, *J. Cell Biol.* **102**, 1284 (1986).
28. D. M. Carrington, A. Auffret, D. E. Hanke, *Nature* **313**, 64 (1985).
29. H. Ito, Y. Fukuda, K. Murata, and A. Kimura, *J. Bacteriol.* **153**, 163 (1983).
30. F. Sherman, G. R. Fink, and J. B. Hicks, *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1983).
31. C. K. Raymond, P. J. O'Hara, G. Eichinger, J. H. Rothman, T. H. Stevens, *J. Cell. Biol.* **111**, 877 (1990).
32. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning, A Laboratory Manual*, 2nd ed. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989).
33. S. D. Emr *et al.*, *J. Cell Biol.* **102**, 523 (1983).
34. R. S. Sikorski and P. Hieter, *Genetics* **122**, 19 (1989).
35. T. A. Kunkel, J. D. Roberts, R. A. Zakour, *Methods Enzymol.* **154**, 367 (1987).
36. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
37. C. J. Roberts, G. Pohlig, J. H. Rothman, T. H. Stevens, *J. Cell Biol.* **108**, 1363 (1989).
38. G. A. Reid, *Methods Enzymol.* **97**, 324 (1983).
39. T. H. Stevens, J. H. Rothman, G. S. Payne, R. Schekman, *J. Cell. Biol.* **102**, 1551 (1986).
40. We thank M. Lindorfer and D. McMillen for the amino-terminal sequencing and oligonucleotide synthesis, and J. Gegner and R. Dahlquist for the pCW-ori+ vector. Supported by a postdoctoral fellowship from the National Institutes of Health (PMK), a Public Health Service predoctoral traineeship (CTY), and Public Health Service grant GM38006 and an American Cancer Society Faculty Research Award (THS).

22 June 1990; accepted 1 October 1990