

Amino acid sequence of the plasma membrane ATPase of *Neurospora crassa*: Deduction from genomic and cDNA sequences

(λ gt11/nucleotide sequence/intron/membrane-spanning segment/amino acid sequence homologies)

KARL M. HAGER*, SUZANNE M. MANDALA*, JAMES W. DAVENPORT*, DAVID W. SPEICHER[†],
EDWARD J. BENZ, JR.*[‡], AND CAROLYN W. SLAYMAN*[§]

Departments of *Human Genetics, [†]Pathology, [‡]Internal Medicine, and [§]Physiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510

Communicated by Edward A. Adelberg, July 7, 1986

ABSTRACT The plasma membrane of *Neurospora crassa* contains an electrogenic H⁺-ATPase (EC 3.6.1.35), for which we have isolated and sequenced both genomic and cDNA clones. The ATPase gene is interrupted by four small introns (58–124 base pairs). It encodes a protein of 920 amino acids (M_r , 99,886) possessing as many as eight transmembrane segments. The *Neurospora* ATPase shows significant amino acid sequence homology with the Na⁺,K⁺- and Ca²⁺-transporting ATPases of animal cells, particularly in regions that appear to be involved in ATP binding and hydrolysis.

The plasma membrane ATPase (EC 3.6.1.35) of the fungus *Neurospora crassa* is a 100-kDa polypeptide that is deeply embedded in the lipid bilayer and constitutes 7–10% of the total membrane protein (1). Its physiological function is to pump protons out of the cell, generating an electrochemical H⁺ gradient that in turn drives the H⁺-dependent uptake of amino acids, sugars, and inorganic ions (2).

The *Neurospora* ATPase has been intensively studied by both physiological and biochemical techniques and as a result has become one of the best-known examples of an electrogenic ion pump. Early studies with micropipette electrodes established that this pump generates a very large membrane potential (to 350 mV, inside negative) and a somewhat smaller pH gradient (1.4 units, inside alkaline) for a total electrochemical gradient of \approx 430 mV (2). More recently, current–voltage analysis has defined the stoichiometry of the pump (one H⁺ transported per ATP split) as well as some of the kinetic parameters of the reaction cycle (3). These electrophysiological results have been complemented and extended by direct measurements of ATP-coupled proton translocation, both in isolated plasma membrane vesicles and in proteoliposomes constructed from purified 100-kDa polypeptide (4–6).

In its molecular structure, as well as in aspects of its reaction mechanism, the *Neurospora* H⁺-ATPase (along with related enzymes from *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*) bears a striking resemblance to the (Na⁺,K⁺)-, Ca²⁺-, and (H⁺,K⁺)-ATPases of animal cells (2). In each case, there is a 100-kDa catalytic subunit that is phosphorylated by ATP to form a covalent β -aspartyl phosphate reaction intermediate. Furthermore, experiments with group-specific reagents have revealed certain key amino acid residues that appear to be conserved: an *N*-ethylmaleimide-sensitive cysteine (7), a butanedione- and phenylglyoxal-sensitive arginine (8, 9), and a fluorescein isothiocyanate (FITC)-sensitive lysine (refs. 10 and 11; J. Pardo, personal communication), all protected by ATP or ADP. Thus, it has seemed reasonable to expect that the

portions of the 100-kDa polypeptide involved in ATP binding and phosphorylation may be quite similar from one enzyme to the next, even though the portions that dictate ion specificity and stoichiometry of ion transport must certainly differ.

Gene cloning provides the most straightforward way to determine the complete amino acid sequence of large proteins such as the ion-translocating ATPases and thus to learn the degree to which they are structurally related. This paper describes the cloning and sequencing of the gene for the *Neurospora* H⁺-ATPase, from which the complete amino acid sequence of its 100-kDa polypeptide has been deduced. Homologies between the *Neurospora* ATPase and other members of the ion-translocating ATPase group are discussed.

METHODS

Preparation of Antibody Specific for the Plasma Membrane ATPase. Polyclonal rabbit antiserum was raised against ATPase purified by the method of Bowman *et al.* (1) followed by preparative NaDodSO₄/polyacrylamide gel electrophoresis. The antiserum was subjected to affinity chromatography on a column of ATPase coupled to Sepharose CL-4B. Upon immunoblotting of *Neurospora* plasma membranes, the affinity-purified antibody specifically recognized a 100-kDa protein that comigrated with the ATPase. In addition, it specifically precipitated the 100-kDa band from *Neurospora* plasma membranes labeled with ³⁵SO₄²⁻.

Screening of cDNA and Genomic Libraries. The affinity-purified ATPase antibody was used to screen a *Neurospora* cDNA library (ref. 12; kindly provided by M. Sachs and U. RajBhandary) in the expression vector λ gt11 (13). Phage DNA was prepared (14) from isolates that were positive through five rounds of screening, and cDNA inserts were subcloned in the plasmid vector pSP65 (15). One clone, cDNA 18, was labeled by nick-translation (16) and used to identify genomic clones within an EMBL4 (17) *Neurospora* library (generously supplied by R. Geever and N. Giles) by plaque hybridization (18). Additional cDNA clones were obtained by screening the λ gt11 cDNA library with nick-translated genomic clone pKH14 (Fig. 1).

DNA Sequencing. The DNA sequence of the genomic and cDNA clones was determined by the dideoxynucleotide-chain termination method of Sanger *et al.* (19) modified for the use of ³⁵S-labeled deoxyadenosine 5'-[α -thio]triphosphate (20). Templates were prepared by inserting restriction fragments into phage M13mp18 or mp19 (21) or by the deletion method of Dale *et al.* (22).

Tryptic Peptide Isolation and NH₂-Terminal Sequencing. For peptide sequencing pure 100-kDa polypeptide was pre-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: bp, base pair(s); FITC, fluorescein isothiocyanate; kb, kilobase(s).

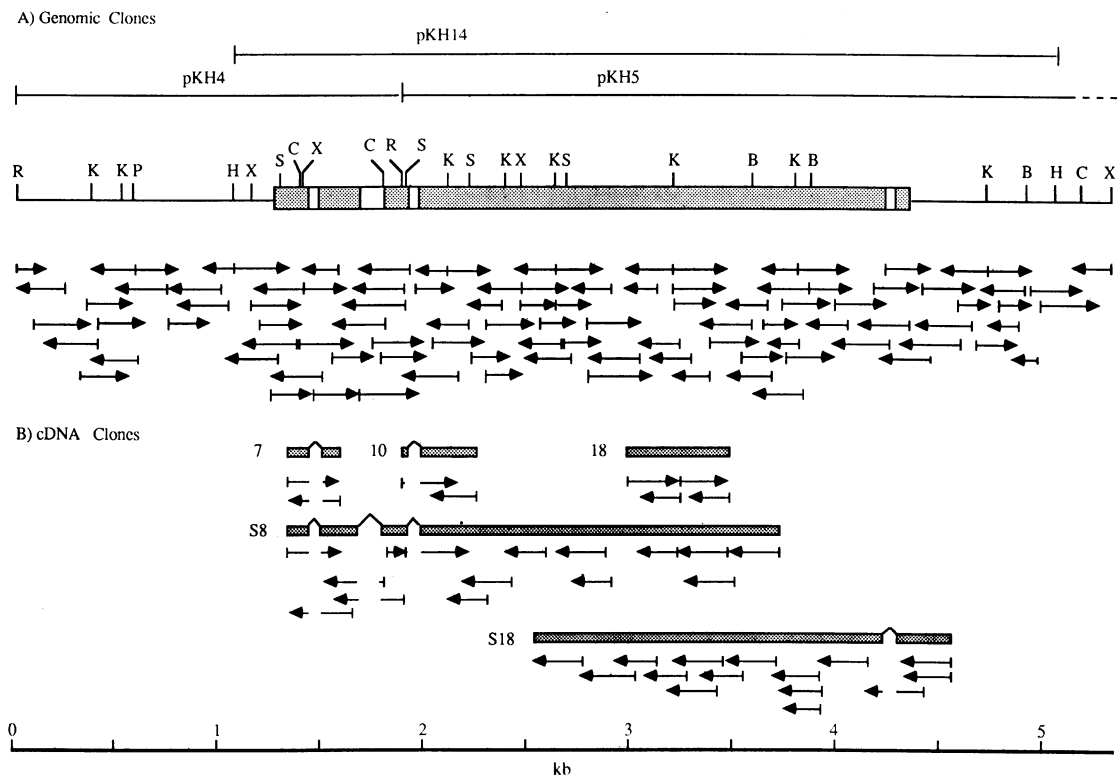


FIG. 1. Structure of the *Neurospora* plasma membrane H^+ -ATPase gene. (A) Restriction map and sequencing strategy for genomic clones pKH4, pKH5, and pKH14. Only the six-base restriction sites *Bgl* II (B), *Cla* I (C), *Hind*III (H), *Kpn* I (K), *Pst* I (P), *Eco*RI (R), *Sal* I (S), *Xho* I (X) used in sequencing are shown. The extent and direction of sequencing are indicated by arrows. Exons are represented by shaded boxes; introns by open boxes. The direction of transcription is from left to right. (B) Structure of the cDNA clones; \blacktriangle represents the introns not present in the cDNA clones. The DNA sequence of the cDNA clones is identical with that of the genomic clones except for a 43-bp sequence (not shown) on the 5' end of clone S18, which presumably is an artifact of cDNA synthesis.

pared by chromatography of glycerol-gradient-purified (1), $NaDodSO_4$ -solubilized ATPase on a Sephacryl S-400 column, followed by extensive dialysis at 4°C to remove $NaDodSO_4$. The intact 100-kDa polypeptide was digested with trypsin (ATPase/trypsin = 50/1, wt/wt) at 37°C for 24 hr, and the resulting peptides were separated by reverse-phase HPLC in 0.1% trifluoroacetic acid, using acetonitrile gradients of 0–100% on a Supelco (Bellefonte, PA) LC-304 column (4.6 mm \times 25 cm). Amino acid analysis and automated Edman degradation were performed as described (23).

RESULTS

Isolation of the *Neurospora* Plasma Membrane ATPase Gene. The first step in isolation of the ATPase gene was to probe a *Neurospora* λ gt11 cDNA library with antibody directed against the purified 100-kDa protein. Of 10^6 plaques screened, 17 gave positive signals through five rounds of plating. Twelve of the 17 plaques yielded small but detectable inserts upon *Eco*RI digestion of purified phage DNA, and the inserts fell into three nonoverlapping classes on the basis of size [193, 312, and 498 base pairs (bp)] and cross-hybridization data (Fig. 1). Representative clones (cDNA 7, 10, and 18) were chosen for further study. DNA from all three clones hybridized to a species of *Neurospora* poly(A)⁺ RNA large enough [4.3 kilobases (kb)] to encode a 100-kDa polypeptide. Furthermore, when RNA homologous to each of the DNA inserts was purified from a preparation of total *Neurospora* poly(A)⁺ RNA by hybrid selection (24), *in vitro* translation in a rabbit reticulocyte lysate generated a 100-kDa polypeptide that comigrated with authentic ATPase (data not shown). Thus, the three cDNA clones appeared to encode different parts of the plasma membrane ATPase.

To obtain the entire ATPase gene, cDNA clone 18 (Fig. 1) was used to screen a *Neurospora* genomic library, yielding a series of large (15- to 20-kb) overlapping genomic clones from which smaller restriction fragments hybridizing with cDNA 7, 10, and 18 were identified and subcloned. Finally, to isolate large cDNA clones, the λ gt11 cDNA library was reprobbed with nick-translated genomic clone pKH14 (Fig. 1). In contrast to the initial antibody screening of the library, which gave cDNA clones with relatively small inserts, clones with inserts up to 1.8 kb were obtained. It is not clear whether the failure to detect large inserts by antibody screening was due to a lack of synthesis of some of the corresponding fusion proteins (for example, because of incorrect reading frames) or to poor detection of relatively large fusion proteins by immunoscreening of plaques.

Nucleotide Sequence of the Plasma Membrane ATPase Gene. Fig. 1 summarizes the strategy employed to sequence both the genomic and cDNA clones, and Fig. 2 illustrates the sequence of the ATPase gene and the amino acid sequence deduced from it. The most probable initiation codon begins at base 1247; it is present in a sequence with homology to the consensus eukaryotic translation start site CCRCCATGG (25) and is the only in-frame methionine codon upstream of the first intron. The ATPase coding region is divided by four small introns, ranging in size from 59 to 124 bp. All four introns contain three highly conserved elements, including the internal sequence RCTRAC that is probably involved in lariat formation during splicing (Fig. 3; ref. 26).

The exons encode a protein of 920 amino acids with a calculated molecular weight of 99,886, a value comparing favorably with a molecular weight of 104,000 (1) determined by $NaDodSO_4$ /polyacrylamide gel electrophoresis. Two lines of evidence support the correctness of the deduced amino acid sequence: (i) the predicted amino acid composi-

Intron	Size	Sequence
1	58 bp	GTAAGT-----35-----TACTAACC-----7-----TAG
2	124 bp	GTACGT-----97-----TGCTGACT-----10-----CAG
3	64 bp	GTAGGT-----36-----CGCTAACC-----11-----CAG
4	67 bp	GTAAGT-----38-----AGCTAACC-----12-----CAG

<i>Neurospora</i> Consensus	GTA ^A GT _C ---35-291---TG ^C CT ^G ACX _A ---6-20---C _T AG
Yeast Consensus	GTATGT-----TACTAACA---5-53---C _T AG

FIG. 3. Introns of the *Neurospora* ATPase gene.

animal cell ATPases (10, 11, 32), appear to contribute to the ATP binding site—one containing an FITC-reactive lysine that can be protected by ATP (residues 473–503) and another containing an 5'-(*p*-fluorosulfonyl)benzoyladenosine-reactive lysine, which also can be protected by ATP (residues 604–663). There is a smaller, though still significant, degree of homology with the more distantly related K⁺-transporting ATPase of *Escherichia coli* (31).

Predicted Transmembrane Orientation. Although the detailed structure of only a few membrane proteins is known, it has proven helpful to use sequential search procedures to identify possible membrane-spanning segments on the basis of their hydrophobicity. In Fig. 5A, the procedure of Engelman *et al.* (33) has been applied to the *Neurospora* H⁺-ATPase. If one takes 20 kcal/mol (1 cal = 4.18J) as a threshold value for the free energy of transfer of a 20-amino acid segment to water, above which value the segment is likely to be located in the membrane (33), the *Neurospora* ATPase is predicted to contain eight transmembrane segments, 20 to 31 residues in length (Fig. 5B). Four of the segments are found in the NH₂-terminal one-third of the protein and four in the COOH-terminal one-third. An additional stretch of 19 amino acids from residue 657 to 675 has a free energy of transfer that is only slightly smaller (18 kcal/mol) and might be regarded as a candidate for a ninth transmembrane segment. This would change the predicted location of the COOH terminus from the cytoplasmic side of the membrane to the extracellular side, a matter that will have

to be resolved by direct biochemical experimentation. The yeast, Na⁺,K⁺- and Ca²⁺-ATPases have strikingly similar hydrophobicity profiles (data not shown) and have been postulated to contain six to ten transmembrane segments, also clustered in the NH₂- and COOH-terminal regions (27, 29, 30, 34, 35).

DISCUSSION

This paper describes the cloning of the gene for the *Neurospora* plasma membrane H⁺-ATPase. The gene contains four introns that are distributed nonrandomly, three clustered at the 5' end and the fourth located quite close to the 3' end; the position of the introns has no obvious relationship to the domain structure of the protein. Of interest, the corresponding yeast gene, though conspicuously homologous in its coding regions, contains no introns (27). In a separate series of experiments not reported here, the *Neurospora* ATPase gene has been mapped by means of a restriction fragment length polymorphism technique to the left arm of chromosome I, close to the *mt* locus (B. J. Bowman and R. L. Metzberg, personal communication).

From the sequencing of both genomic and cDNA clones, the primary structure of the *Neurospora* H⁺-ATPase has been deduced and a working model of its topology in the membrane has been developed. The model predicts that approximately 22% of the protein is embedded in the membrane, 74% protrudes into the cytoplasm, and only about 4% is exposed at the outer surface. The two major cytoplasmic loops show significant homology to the corresponding portions of the Na⁺,K⁺- and Ca²⁺-ATPases, suggesting that they participate in common functions such as ATP binding and hydrolysis. The availability of cloned genes for several members of the ion-translocating ATPase family will pave the way for a direct test of this idea and may also permit the identification of sequences involved in ion specificity.

We thank S. Baserga, L. Caballero, E. Feingold, F. Kalousck, J. Kraus, T. Lalor, R. Mercer, J. Metherall, J. Rose, and J. Schneider for technical advice and helpful discussions; M. Sachs, U. RajBhandary, R. Geever, and N. Giles for providing *Neurospora*

H	191	T[E]A[F]E[V]V[P]G[D]I[L]Q[V]E[E]G[T]I[F]A[D]G[R]I[V]T[D]D . A[F]L[Q]V[D]Q[S]A[L]T[G]E[S]L[A]V[D]K[H]	240
Ca	140	I[K]A[R]D[I]V[P]G[D]I[V]E[V]A[V]G[D]K[V]P[A]D[I]R[L]S[I]K[S]T[L]R[V]D[Q]S[I]L[T]G[E]S[V]S[V]I[K]H	190
Na, K	178	I[N]A[E]E[V]V[V]G[D]L[V]E[V]K[G]G[D]R[I]P[A]D[L]R[I]S[A]N . . GCKVNDNSLITGESEFQTRS	226
Kdp	121	V[P]A[D]Q[L]R[K]G[D]I[V]L[V]E[A]G[D]I[F]P[C]D[G]E[V]L[E]G G . . . A[S]V[D]E[S]A[T]T[G]E[S]A[P]V[I]R[E]	168
H	245	V[F]A[S]S[A]V[K]R[G]E[A]F[V]V[I]T[A]T[G]D[N]T[F]V[G]R[A]A[A]L[V]N[A]A[S]G[G]S[G]H	285
Ca	208	L[F]S[G]T[N]I[A]A[G]K[A]L[G]I[V]A[T]T[G]V[S]T[E]I[G]K[I]R[D]Q[M]A[A]T[E]Q[D]K[T]P	248
Na, K	241	A[F]F[S]T[N]C[V]E[G]T[A]R[G]I[V]V[Y]T[G]D[R]T[V]M[G]R[I]A[T]L[A]S[G]L[E]G[G]Q[T]P	281
H	327	L[A]I[T]I[G]V[F]V[G]L[P]A[V]V[T]T[M]A[V]G[A]Y[L]A[K]K[K]A[I]V[Q]K[L]S[A]I[F]S[L]A[G]V[E]I[L]C[S]D[K]T[G]T[L]K[N]	386
Ca	300	V[A]L[A]V[A]A[I]P[E]G[L]P[A]V[I]T[T]C[L]A[L]G[T]R[M]A[K]K[N]A[I]V[R]S[L]P[S]V[E]T[L]G[C]T[S]V[I]C[S]D[K]T[G]T[L]T[N]	359
Na, K	323	I[G]I[V]A[N]V[P]E[G]L[L]A[T]V[T]V[C]L[F]L[A]T[K]R[M]A[R]K[N]C[L]V[K]N[L]E[A]V[E]T[L]G[S]T[S]T[I]C[S]D[K]T[G]T[L]T[N]	382
Kdp	256	V[A]L[L]V[C]L[I]P[T]T[I]G[G]L[S]A[S]A[V]A[G]M[S]R[M]L[G]A[N]V[I]A[T]S[G]R[A]V[E]A[A]G[D]V[D]V[L]L[D]K[T]G[T]T[L]G	315
H	471	T[C]V[K]G[A]P[L]F[V]L[K]T[V]E[E]D[H]P[I]P[E]V[D]Q[A]Y[K]N[K]V[A]E[F]	505
Ca	512	M[F]V[K]G[A]P[E]G[V]I[D]R[C]N[Y]V[R]V[G]T[T]R[V]P[M]T[G]P[V]K[E]K[I]L	546
Na, K	503	L[V]M[K]G[A]P[E]R[I]L[D]R[C]S[S]I[L]I[H]G[K]E[Q]P[L]D[E]E[L]K[D]A[F]Q	537
Kdp	392	M[I]R[K]G[S]V[D]A[I]R[R]H[V]E[A]N[G]G[H]F[P]T[D]V[D]Q[K]V[D]Q[V]A[R]Q	426
H	534	D[P]P[R]H[D]T[Y]K[T]V[C]E[A]K[T]L[G]L[S]I[K]M[L]T[G]D[A]V[G]I[A]R	566
Ca	601	D[P]P[R]K[E]V[M]G[S]I[Q]L[C]R[D]A[G]I[R]V[I]M[I]T[G]D[N]K[G]T[A]I	633
Na, K	591	D[P]P[R]A[A]V[P]D[A]V[G]K[C]R[S]A[G]I[K]V[I]M[V]T[G]D[H]P[I]T[A]K	623
Kdp	447	D[I]V[K]G[G]I[K]E[A]F[A]Q[L]R[K]M[G]I[K]T[V]M[I]T[G]D[N]R[L]T[A]A	479
H	604	A[D]G[F]A[E]V[F]P[Q]H[K]Y[N]V[E]I[L]Q[R]G[Y]L[V]A[M]T[G]D[G]V[N]D[A]F[S]L[K]K[A]D[T]G[I]A[V]E[G] . S[S]D[A]A[R]S[A]A[D]	663
Ca	673	A[C]C[F]A[R]V[E]P[S]H[K]S[K]I[V]E[Y]L[Q]S[Y]D[E]I[T]A[M]T[G]D[G]V[N]D[A]P[A]L[K]K[A]E[I]G[I]A[M]G[S] . G[T]A[V]A[K]T[A]S[E]	732
Na, K	685	E[I]V[F]A[R]T[S]P[Q]Q[K]L[I]I[V]E[G]C[Q]R[G]A[I]V[A]V[T]G]D[G]V[N]D[S]P[A]L[K]K[A]D[I]G[V]A[M]G[I]A[G]S[D]V[S]K[Q]A[A]D	745
Kdp	488	D[D]F[L]A[E]A[T]P[E]A[K]L[A]L[I]R[Q]Y[Q]A[E]G[R]L[V]A[M]T[G]D[G]T[N]D[A]P[A]L[A]Q[A]D[V]A[V]A[M]N[S] . G[T]Q[A]A[K]E[A]G[N]	547

FIG. 4. Amino acid homology between the *Neurospora* plasma membrane H⁺-ATPase and other ion-translocating ATPases. Homologies were identified using the University of Wisconsin Genetics Computer Group programs (28), introducing gaps (indicated by ●) to maximize homology. Identical residues are boxed; the phosphorylation site is indicated by * and the FITC site by ↓. Designation of sequences is as follows: H, the *Neurospora* plasma membrane H⁺-ATPase sequenced in this work; Ca, the fast-twitch form of the rabbit sarcoplasmic reticulum Ca²⁺-ATPase (29); Na,K, the sheep kidney Na⁺,K⁺-ATPase (30); and Kdp, the product of the *E. coli* *KdpB* gene (31). The H⁺-ATPase of yeast shows similar sequence homology with the animal cell and *E. coli* ion-translocating ATPases (27).

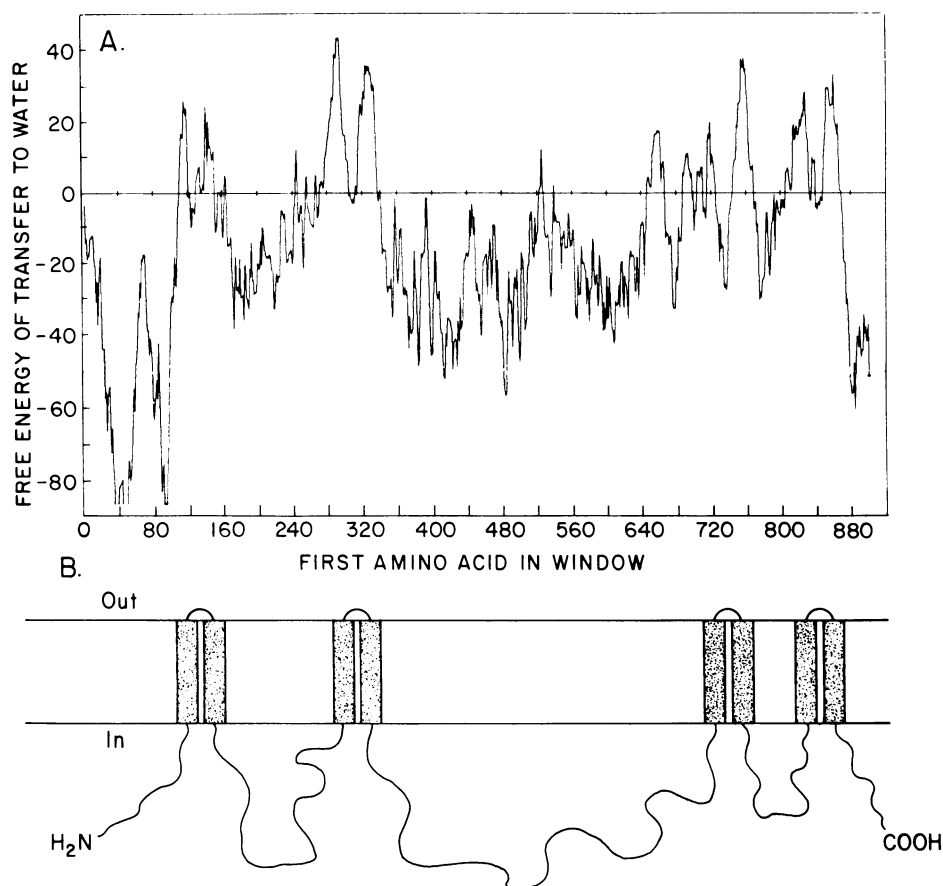


FIG. 5. (A) Hydropathy plot of the *Neurospora* plasma membrane ATPase, generated by the method of Engelman *et al.* (33) using a window of 20 amino acids. The free energy of transfer from a lipid phase to water (in kcal/mol of 20-amino acid segment) is plotted as a function of the first amino acid in the window. Hydrophobic regions are above the x-axis and hydrophilic regions below it. (B) Model of the transmembrane orientation of the *Neurospora* plasma membrane ATPase. In and Out refer to the cytoplasmic and external sides of the membrane, respectively.

libraries; M. Hart and D. Engelman for conducting the hydropathy analysis; and R. Serrano and G. Fink for making available a preprint of the yeast ATPase work. This study was supported by National Institutes of Health Grant GM 15761. K.M.H. was the recipient of a National Institutes of Health Predoctoral Traineeship (GM 07499); S.M.M., of a National Science Foundation Postdoctoral Fellowship (DMB 8508818); and J.W.D., of a National Institutes of Health Postdoctoral Fellowship (GM 10363).

1. Bowman, B. J., Blasco, F. & Slayman, C. W. (1981) *J. Biol. Chem.* **256**, 12343-12349.
2. Goffeau, A. & Slayman, C. W. (1981) *Biochim. Biophys. Acta* **639**, 197-223.
3. Slayman, C. W. & Sanders, D. (1984) in *Hydrogen Ion Transport in Epithelia*, eds. Forte, J. G., Warnock, D. G. & Rector, F. C., Jr. (Wiley, New York), pp. 47-56.
4. Perlin, D. S., Kasamo, K., Brooker, R. J. & Slayman, C. W. (1984) *J. Biol. Chem.* **259**, 7884-7892.
5. Scarborough, G. A. & Addison, R. (1984) *J. Biol. Chem.* **259**, 9109-9114.
6. Perlin, D. S., San Francisco, M. J. D., Slayman, C. W. & Rosen, B. P. (1986) *Arch. Biochem. Biophys.* **248**, 53-61.
7. Brooker, R. J. & Slayman, C. W. (1983) *J. Biol. Chem.* **258**, 222-226.
8. DiPietro, A. & Goffeau, A. (1985) *Eur. J. Biochem.* **148**, 35-39.
9. Kasher, J., Allen, K. E., Kasamo, K. & Slayman, C. W. (1986) *J. Biol. Chem.* **261**, 10808-10813.
10. Farley, R. A., Tran, C. M., Carilli, C. T., Hawke, D. & Shively, J. E. (1984) *J. Biol. Chem.* **259**, 9532-9535.
11. Mitchinson, C., Wilderspin, A. F., Trinnaman, B. J. & Green, N. M. (1982) *FEBS Lett.* **146**, 87-92.
12. Sachs, M. S., David, M., Werner, S. & RajBhandary, U. L. (1986) *J. Biol. Chem.* **261**, 869-873.
13. Young, R. A. & Davis, R. W. (1983) *Science* **222**, 778-782.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 76-85.
15. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984) *Nucleic Acids Res.* **12**, 7035-7070.
16. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237-251.
17. Frischauf, A. M., Lehrach, H., Poustka, A. & Murray, N. (1983) *J. Mol. Biol.* **170**, 827-842.
18. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180-182.
19. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
20. Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3963-3965.
21. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33**, 103-119.
22. Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) *Plasmid* **13**, 31-40.
23. Speicher, D. W., Davis, G., Yurchenco, P. D. & Marchesi, V. T. (1983) *J. Biol. Chem.* **258**, 14931-14937.
24. Parnes, J. R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, V., Appella, E. & Seidman, J. G. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2253-2257.
25. Kozak, M. (1984) *Nucleic Acids Res.* **12**, 857-872.
26. Rodriguez, J. R., Pikielny, C. W. & Rosbash, M. (1984) *Cell* **39**, 603-610.
27. Serrano, R., Kielland-Brandt, M. C. & Fink, G. R. (1986) *Nature (London)* **319**, 689-693.
28. Devereux, J. R., Haekerli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387-395.
29. Brandl, C. J., Green, N. M., Korczak, B. & MacLennan, D. H. (1986) *Cell* **44**, 597-607.
30. Schull, G. E., Schwartz, A. & Lingrel, J. B. (1985) *Nature (London)* **316**, 691-695.
31. Hesse, J. E., Wiczorek, L., Altendorf, K., Reicin, A. S., Dorus, E. & Epstein, W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4746-4750.
32. Ohta, T., Nagano, K. & Yoshida, M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2071-2075.
33. Engelman, D. M., Steitz, T. A. & Goldman, A. (1986) *Annu. Rev. Biophys. Chem.* **15**, 321-353.
34. Kawakami, K., Noguchi, S., Noda, M., Takahashi, H., Ohta, T., Kawamura, M., Nojima, H., Nagano, K., Hirose, T., Inayama, S., Hayashida, H., Miyata, T. & Numa, S. (1985) *Nature (London)* **316**, 733-736.
35. MacLennan, D. H., Brandl, C. J., Korczak, B. & Green, N. M. (1986) *Nature (London)* **316**, 696-700.