ELSEVIER

Biochimica et Biophysica Acta 1319 (1997) 9–13

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

Short sequence-paper

Cloning and sequencing of the gene encoding the plasma membrane H^+-ATPase from an acidophilic red alga, *Cyanidium caldarium* [1]

Hisataka Ohta [a], Hitoshi Shirakawa [b], Kohji Uchida [c], Michiteru Yoshida [b], Yuhsi Matuo [c], Isao Enami [a,*]

[a] *Department of Biology, Faculty of Science, Science University of Tokyo, Kagurazaka, Shinjuku-ku, Tokyo 162, Japan*
[b] *Department of Biological Science and Technology, Faculty of Industrial Science and Technology, Science University of Tokyo, Yamazaki, Noda, Chiba 278, Japan*
[c] *Nagahama Institute for Biochemical Science, Oriental Yeast Co. Ltd., Nagahama, Shiga 526, Japan*

Received 12 November 1996; revised 17 January 1997; accepted 21 January 1997

Abstract

A cDNA containing an open reading frame encoding the putative plasma membrane H^+-ATPase in an acidophilic red alga, *Cyanidium caldarium*, was cloned and sequenced by means of PCR and Southern hybridization based on homologous sequences of P-type ATPases found in other organisms. The cloned cDNA is 3300 bp in length, containing a 2865 bp open reading frame encoding a polypeptide of 955 amino acids which has a predicted molecular mass of 105 371. The deduced amino acid sequence was found to be more homologous to those of P-type H^+-ATPases from higher plants than that from the green alga *Dunaliella bioculata*.

Keywords: P-type H^+-ATPase; Intracellular pH regulation; cDNA sequence; Acidophilic red alga; (*Cyanidium caldarium*)

Cyanidium caldarium is an acidophilic and thermophilic unicellular red alga found in acidic hot springs. This alga is the only photosynthetic organism which can grow at pHs as low as 0 [1]. The strain examined in this study is an obligate photoautotroph, and has a discrete nucleus, mitochondria and a large single chloroplast but seems to contain no vacuoles as observed under electron microscopy [2,3]. The intracellular pH of the *Cyanidium* cells is known to be neutral, in spite of growing at extremely acidic pHs [4–6]. Enami et al. have directly determined the intracellular pH by ^{31}P nuclear magnetic resonance and found that the internal pH of the *Cyanidium* cells is maintained in the neutral pH region by an active H^+ efflux (H^+-pump) mediated by a plasma membrane ATPase [6]. The activity of this H^+-pump depends on the cellular metabolic activities (photosynthesis or respiration) [7–9]. Since the alga has no vacuole, the plasma membrane ATPase is considered to be the major source in regulating the intracellular pH of the red alga. It should be pointed out here that, although *Cyanidium caldarium* having vacuole was also reported [10], it is a different strain from our one and recently it has been renamed *Galdieria sulphuraria* [3]. In this work, the cDNA encoding the putative plasma membrane H^+-ATPase of *Cyanidium*

* Corresponding author. Fax: +81 471 24-2150; E-mail: enami@rs.noda.sut.ac.jp

[1] The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases with the following accession number D88424.

0005-2728/97/$17.00
PII S0005-2728(97)00007-8

```
AAACCGTCTTTCTTTGGCCAGTGTCTGCGAGTGCGCTTTCATCGCTATCCGTGTCAGCGCACGAGCCAGTTTTGTATTGGAGGCTTTTTCTGCGGCACTGGGACGACGTTTGAACGTACA  120
ACGATAGCTGTACTTGGGCAGCGCTGGAACGTTTCGCGGTAAACTCGAGACGTACGCGTCGCTTCTTTCTATCTTAGCAAGGCACGGAAACGAAGTGCCGAACAGCAAATCATTATGGCG  240
                                                                                                                  M  A
GACGTCGAGTCTGGAAACTCGCAGGTGAATGGAGGGAATGAGGTGGACGGCACGAACGGTTCATCTAGTTTGAGGGCGGAGTCTAAACCGTTACCCGCCGACTTGGATAAGCGTTCGTTG  360
D  V  E  S  G  N  S  Q  V  N  G  G  N  E  V  D  G  T  N  G  S  S  S  L  R  A  E  S  K  P  L  P  A  D  L  D  K  R  S  L
TCTATGAGCATGCTGGAAACGGTAGACCTTGAGAAGGATGACATGGACTATGTGATGGCGTGTTTGAAGACGTCGCCCGAGGGGTTGAAGCCGGATGTTGCAGCGCGTCGTTTAGCAAAG  480
S  M  S  M  L  E  T  V  D  L  E  K  D  D  M  D  Y  V  M  A  C  L  K  T  S  P  E  G  L  K  P  D  V  A  A  R  R  L  A  K
TTTGGTCCGAATGCGCTACCCGAGAAGAAGGTTAACCCGATTCTGGAGTTCTTGATGTTTATGTGGAATCCTCTGAGTTGGGTGATGGAGGCAGCTGCTTTGGTTGCCATCTTTTTGACA  600
F  G  P  N  A  L  P  E  K  K  V  N  P  I  L  E  F  L  M  F  M  W  N  P  L  S  W  V  M  E  A  A  A  L  V  A  I  F  L  T
ATTCCTGGTGGGAAAGCTCCTGATTGGGAAGATTTTCTTGGCATTCTCTTGCTGCTTTTGATCAACTCCACGATCGGTTTCATTGAGGAACGCAATGCTGGCAATGCTGTGAAAGCACTT  720
I  P  G  G  K  A  P  D  W  E  D  F  L  G  I  L  L  L  L  L  I  N  S  T  I  G  F  I  E  E  R  N  A  G  N  A  V  K  A  L
ATGGATGCACTGGCGCCCCGGGCCAAGGTGCAAAGGGGTGGTGAATGGCTCGATATCGATGCGAAAGATCTCGTCATCGGTGATATTGTTGCGTTGAAACTGGGCGATGTGATCCCTGCT  840
M  D  A  L  A  P  R  A  K  V  Q  R  G  G  E  W  L  D  I  D  A  K  D  L  V  I  G  D  I  V  A  L  K  L  G  D  V  I  P  A
GATGCGCGGATCATGAACGGCAAGGATATCAAGATTGACCAAGCAGCGCTCACCGGTGAATCGCTTCCTGTTGGAAAAGAAAAAGGTGATATGATCTATTCAGGATCTGTGGTGAAACAG  960
D  A  R  I  M  N  G  K  D  I  K  I  D  Q  A  A  L  T  G  E  S  L  P  V  G  K  E  K  G  D  M  I  Y  S  G  S  V  V  K  Q
GGAGAGTTCCTGGCGCTTGTGATTGCCACCGGCATGAACACGTTCTTTGGCAAAGCTGCTCACCTGGTGAATCAGACAGAGAGCACTTCACATTTGCAGGCGATTGTCTCTGCGATCGGC 1080
G  E  F  L  A  L  V  I  A  T  G  M  N  T  F  F  G  K  A  A  H  L  V  N  Q  T  E  S  T  S  H  L  Q  A  I  V  S  A  I  G
CTCTACTGTATGGCCTGGATCAGCACCTTTGTTCTTCTCTTGATTGTAACGCAGTGGCCAATACATTTGGAGAACTATCGTCATGGAATCAACAACATCCTGGTGCTCTTGATCGGTGGT 1200
L  Y  C  M  A  W  I  S  T  F  V  L  L  L  I  V  T  Q  W  P  I  H  L  E  N  Y  R  H  G  I  N  N  I  L  V  L  L  I  G  G
GTCCCCATCGCGATGCCTGTTGTGCTCTCGGTCACACTGGCTATCGGTGCGCACGAACTTGCCGAACAAAAAGCGATCGTCACCCGGATGACTGCCGTCGAAGAACTCGCAGGTATGACC 1320
V  P  I  A  M  P  V  V  L  S  V  T  L  A  I  G  A  H  E  L  A  E  Q  K  A  I  V  T  R  M  T  A  V  E  E  L  A  G  M  T
ATCCTCTGCTCGGATAAGACTGGTACGCTGACCCTCAACAAACTTTCGATTGACCAAGAAAGCTTCTTTACCATGGGCGGCTACACCGTCGATACCGTCGACCAGTGCATGGTCTTCGCT 1440
I  L  C  S  D  K  T  G  T  L  T  L  N  K  L  S  I  D  Q  E  S  F  F  T  M  G  G  Y  T  V  D  T  V  D  Q  C  M  V  F  A
GCTCGTGCGTCTCGCACAGAGAACCAAGATGCCATCGACTTTGCGGTGGTCAACTCTCTACCAGATCCCAAGATGGCCCGCGAAGGCATTGAAGAGCTGGACTTTCATCCGTTCAACCCG 1560
A  R  A  S  R  T  E  N  Q  D  A  I  D  F  A  V  V  N  S  L  P  D  P  K  M  A  R  E  G  I  E  E  L  D  F  H  P  F  N  P
GTTGATAAGCGCACAGAAATCACCTACCGCGACAACAAAGATGGCAAAGTCTACAAAGCAACCAAAGGGGCACCCCAGATCATCCTTGGCATGGCCCACAACAAGAAGGAAATTGAAAAA 1680
V  D  K  R  T  E  I  T  Y  R  D  N  K  D  G  K  V  Y  K  A  T  K  G  A  P  Q  I  I  L  G  M  A  H  N  K  K  E  I  E  K
GAAGTTCATGAACAGATTGAAGATTTCGCGAAACGCGGCTTCCGTGCACTTGGCATTGCAGTCGCGGAGGTGCCGTCCGGAGAGGCGCACGGCGAACCCGGCCCCTGGTCTATGGTCGGT 1800
E  V  H  E  Q  I  E  D  F  A  K  R  G  F  R  A  L  G  I  A  V  A  E  V  P  S  G  E  A  H  G  E  P  G  P  W  S  M  V  G
TTGATGCCTATTTTCGACCCGCCGCGTCACGACACGAAGGAGACGATCGAGCAAGCCATCGCCATGGGTGTCGAAGTGAAGATGATCACCGGTGACCAGCTGGCGATCGCCAAGGAAACC 1920
L  M  P  I  F  D  P  P  R  H  D  T  K  E  T  I  E  Q  A  I  A  M  G  V  E  V  K  M  I  T  G  D  Q  L  A  I  A  K  E  T
GCTCGTCGTCTTGGAATGGGCACGAATATCTTCAACACGGATGTGCTGAATTTGAGCGACCAGCGGGCCTCCATCGAGTACGGTGGCAGTGTCGGTGAGTTGGTGGAAAGTGCGGATGGT 2040
A  R  R  L  G  M  G  T  N  I  F  N  T  D  V  L  N  L  S  D  Q  R  A  S  I  E  Y  G  G  S  V  G  E  L  V  E  S  A  D  G
TTCGCTGGCGTCTTTCCGGAGCACAAATATCGCATCGTGGAGGTGCTTCAACGGCGTGGCCATATGGTCGGCATGACGGGTGACGGTGTGAACGATGCGCCCGCGCTGAAGAGAGCATCT 2160
F  A  G  V  F  P  E  H  K  Y  R  I  V  E  V  L  Q  R  R  G  H  M  V  G  M  T  G  D  G  V  N  D  A  P  A  L  K  R  A  S
GTGGGCATTGCGGTGGCTGGCGCCACCGACGCAGCCCGTGGAGCCTCAGACATTGTTTTAACAGAACCAGGTCTGAGCGTCATCATTCATGCGATGGTCATGAGCCGTCAGATTTTCCAG 2280
V  G  I  A  V  A  G  A  T  D  A  A  R  G  A  S  D  I  V  L  T  E  P  G  L  S  V  I  I  H  A  M  V  M  S  R  Q  I  F  Q
CGTATGAAGAACTATTCCATGTACGCTTGCTCAGTGACGGTCCGTATTGTGGTGACTTTTTCGATTCTGGTCTGGGCGTTCCGCTTCAACATGCCACCTTTTCTGGTGCTGATTCTTGCG 2400
R  M  K  N  Y  S  M  Y  A  C  S  V  T  V  R  I  V  V  T  F  S  I  L  V  W  A  F  R  F  N  M  P  P  F  L  V  L  I  L  A
TATCTGAATGACGGCACAATCATGACAATTAGCAAGGATCGAGTGAAACCGAGTCCACTGCCGCAGCGATGGGATCTGAAGGAGGTCTTTATTGTGGCCTCCTCATTGGGTATCTACCTG 2520
Y  L  N  D  G  T  I  M  T  I  S  K  D  R  V  K  P  S  P  L  P  Q  R  W  D  L  K  E  V  F  I  V  A  S  S  L  G  I  Y  L
ACGGCAAGCACGGTGATCTTCTACGTGACCCTCTTCAAGACACAGTTCTGGCATGATACGTTCAAGCTAGGGATGCCATGGCTGAACCCTCGAGATCCAAACTACTTCCAGCTGCATTCT 2640
T  A  S  T  V  I  F  Y  V  T  L  F  K  T  Q  F  W  H  D  T  F  K  L  G  M  P  W  L  N  P  R  D  P  N  Y  F  Q  L  H  S
ATCATCTACCTACAGGCTAGCATCATCGGCCAAGCGCTGATTTTCGTCACCCGTGCACACTGGTTCTTCTTTATGGATCGTCCTGGTATCCTGCTCATGAGCGCCTTCGTCGTTGCCCAG 2760
I  I  Y  L  Q  A  S  I  I  G  Q  A  L  I  F  V  T  R  A  H  W  F  F  F  M  D  R  P  G  I  L  L  M  S  A  F  V  V  A  Q
CTTGTCGCTACATTTATCTGCGTCTATGCGAACTGGGGTTTCACTCAGATTCAGGGAACTGGTTGGGGCTGGGCTGGAGTCGTCTGGGTCTGGAACGTCATCTGGTACGCACCTTTGGAC 2880
L  V  A  T  F  I  C  V  Y  A  N  W  G  F  T  Q  I  Q  G  T  G  W  G  W  A  G  V  V  W  V  W  N  V  I  W  Y  A  P  L  D
ATTATTAAGATCGCTGTGCGAAGCATTATCACTGGCGACAAGACCCCAATCCATAAGCTCTTCGCAGCACGGCGCATGTTTACATTCGACTACTCGAAGCATGGTCGCGAAGGTCGCATG 3000
I  I  K  I  A  V  R  S  I  I  T  G  D  K  T  P  I  H  K  L  F  A  A  R  R  M  F  T  F  D  Y  S  K  H  G  R  E  G  R  M
CCGCGCAGCAGCCTCCAGGCAGCTCAAGCGCGCGCGTCTGTGCATCGCTCAATGGAGACATATCGAGCCTCGCTGCAGAAGAATGTGAATTCCCTGGACTAGCACAAGCGGATGCGAGCA 3120
P  R  S  S  L  Q  A  A  Q  A  R  A  S  V  H  R  S  M  E  T  Y  R  A  S  L  Q  K  N  V  N  S  L  D  *
AAGTAGTGATTCCAGACTGGGAAAGGGCCGTGCAACGGACACGGCAACACGCCGCGGTGGAAAAACATAGCGAACAAAAGGTGGTTCATGGCGTCGTCCGACAGATGGCGAGTGTACTAC 3240
TGGCTCAGAGGAACGGTTTCTGAATGTGTAAAACAGCGCAGCATCGTGCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Fig. 1. Nucleotide sequence of the gene encoding the plasma membrane H^+-ATPase from the acidophilic red alga, *Cyanidium caldarium* and the deduced amino acid sequences. The nucleotide sequences used as degenerate oligonucleotide primers are underlined; the putative initiation codon is double underlined; and the putative polyadenylation signal is boxed.

caldarium was cloned and sequenced, and the deduced amino acid sequence was compared with those of known H^+-ATPases from other organisms.

C. caldarium was photoautotrophically grown under continuous illumination in a medium of pH 3 at 40°C, as described previously [4]. For preparation of cDNA and cDNA library of *C. caldarium*, total RNA was prepared from *C. caldarium* and then poly(A)$^+$ RNA was purified from the total RNA by oligo(dT) Sepharose spun column (Pharmacia Biotech). Double-stranded cDNA was prepared from the poly(A)$^+$ RNA using reverse transcriptase with oligo(dT) as a primer, and the library was constructed in the phage λgt22A with a Gigapack II Gold kit from Stratagene. Two amino acid sequences conserved in all P-type ATPases, DKTGTL (phosphorylation site) and GDGVND (ATP-binding site) (underlined sequences in Fig. 1), were chosen to synthesize degenerate oligonucleotide primers, in which *Eco*RI restriction sites were flanked to the respective 5′-end. Using these primers, a cDNA fragment was amplified using the double stranded cDNA as template by polymerase chain reaction (PCR). The resulted 800 bp fragment was recovered from an agarose gel with GENECLEAN (Bio 101), digested with *Eco*RI, and then inserted into the plasmid Bluescript II KS(+) (Stratagene). Four independent clones were purified and sequenced. All of the sequences obtained were identical and were highly homologous to P-type H^+-ATPases known in plants (data not shown). To obtain cDNA clone containing the entire coding sequence, the λgt22A cDNA library of *C. caldarium* was screened by Southern hybridization with the 800 bp PCR fragment as a probe. Inserts from positive phages were cloned into Bluescript II KS(+), and the resulted clone was sequenced from both directions using a Taq Dye Primer Cycle Sequencing Kit (Perkin Elmer).

Fig. 1 shows the nucleotide sequence (about 3300 bp) of the cDNA clone obtained. This sequence

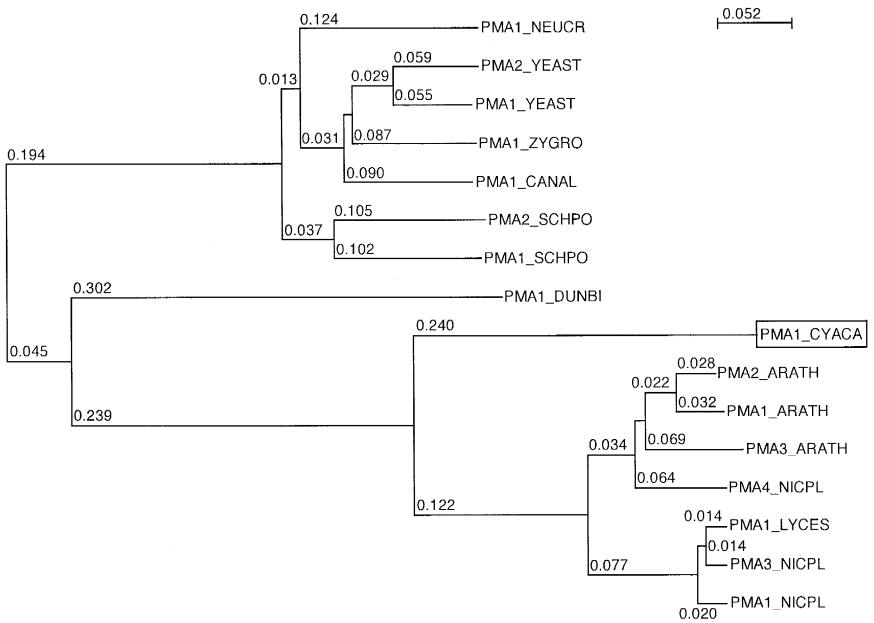

Fig. 2. Phylogenetic analysis of the plasma membrane H^+-ATPase based on the amino acid sequences currently available from various organisms. The cladogram was constructed from 15 different sequences found in the SWISS-PROT database plus the *Cyanidium caldarium* H^+-ATPase sequence reported here, and the homology index was calculated with the program Clastal W [22]. Distances along the horizontal axis are proportional to the difference between sequences. ARATH, *Arabidopsis thaliana*; CANAL, *Candida albicans*; CYACA, *Cyanidium caldarium*; DUNBI, *Dunaliella bioculata*; LYCES, *Lycopersicon esculentum*; NEUCR, *Neurospora crassa*; NICPL, *Nicotiana plumbaginifolia*; SCHPO, *Schizosaccharomyces pombe*; YEAST, *Saccharomyces cerevisiae*; and ZYGRO, *Zygosaccharomyces rouxii*.

contained a putative open reading frame of 2865 bp in length starting from the sequence ATGGC (double underlined in Fig. 1) which is consensus to the start codon for translation initiation in plants [11]. The sequence around the start codon (ATTATGG) also fulfills the requirements for the optimal sequence around the ATG codon for initiation by eukaryotic ribosomes as reported by Kozak [12]. At a site 21 bp upstream from the poly(A) sequence, there is a putative polyadenylation signal sequence (TGTAA) (boxed in Fig. 1) [13]. These features supported our assiginment of the 2865 bp open reading frame. According to this assignment, the open reading frame encodes a protein of 955 amino acid residues with a calculated molecular mass of 105 371. Comparison of the deduced amino acid sequence with those of five P-type ATPases (Ca^{2+}-, Na^+/K^+-, H^+/K^+-, H^+- and K^+-ATPase) (for a review, see [14]) revealed the highest homology to H^+-ATPase (data not shown). Thus, we performed a homology search of the putative H^+-ATPase from *C. caldarium* with the corresponding P-type H^+-ATPase from various sources currently available. A search in the Swiss-Prot database revealed that the *C. caldarium* H^+-ATPase is 55–60% identical to the P-type H^+-ATPase from higher plants (*Arabidopsis thaliana* [15], *Lycopersicon esculentum* [16] and *Nicotiana plumbaginifolia* [17]) but only 36–38% identical to the corresponding proteins from *Neurospora crassa* [18] and *Saccharomyces cerevisiae* [19] (data not shown). Fig. 2 shows a cladogram of all available P-type H^+-ATPase sequences. As expected, H^+-ATPases can be divided into two groups, one found in fungi and the other in plants. So far, only one other algal P-type H^+-ATPase has been described; this is from the salt-tolerant green alga *Dunaliella bioculata* [20]. Surprisingly, the new sequence from *C. caldarium* is more homologous to those of higher plants than to that of *Dunaliella bioculata* (Fig. 2), although a more systematic comparison cannot be made until the sequences of other algal P-type H^+-ATPases become known. Hydropathy analysis of the *C. caldarium* H^+-ATPase predicted eight transmembrane helices (data not shown) which is consistent with the current structural model of fungal and higher plant H^+-ATPases [14]. These results make it likely that the cDNA cloned in the present study is indeed the gene encoding the P-type H^+-ATPase which might be responsible for intracellular pH regulation in the acidophilic unicellular red alga *C. caldarium*. The C-terminal region in the yeast and plant ATPases has been reported to serve as an autoinhibitory domain important for regulation of its activity (for review, see [21]). A homologous sequence corresponding to the autoinhibitory domain, however, was not found in the deduced C-terminal region of the *C. caldarium* H^+-ATPase gene cloned here. Studies on the physiological function of the cloned H^+-ATPase, the possible role of C-terminal region, as well as their relationship with the extremely strong acid-tolerance of the red alga *C. caldarium*, are in progress.

We thank Dr. J.-.R. Shen of the Institute of Physical and Chemical Research (RIKEN) for his critical reading of the manuscript.

References

[1] Allen, M.B. (1959) Arch. Microbiol. 32, 270–277.
[2] Enami, I., Nagashima, H. and Fukuda, I. (1975) Plant Cell Physiol. 16, 221–231.
[3] Seckbach, J. (1994) in Evolutionary pathways and enigmatic algae: *Cyanidium caldarium* [Rhodophyta] and related cells (Seckbach, J., ed.), Kluwer Academic, Dordrecht, pp. 99–132.
[4] Enami, I. and Fukuda, I. (1975) Plant Cell Physiol. 16, 211–220.
[5] Enami, I. (1978) Plant Cell Physiol. 19, 869–876.
[6] Enami, I., Akutsu, H. and Kyogoku, Y. (1986) Plant Cell Physiol. 27, 1351–1359.
[7] Kura-Hotta, M. and Enami, I. (1981) Plant Cell Physiol. 22, 1175–1183.
[8] Enami, I. and Kura-Hotta, M. (1984) Plant Cell Physiol. 25, 1107–1113.
[9] Kura-Hotta, M. and Enami, I. (1984) Plant Cell Physiol. 25, 1115–1122.
[10] Zielgler, K., Hauska, G. and Nelson, N. (1995) Biochim. Biophys. Acta 1230, 202–206.
[11] Lütcke, H.A., Chow, K.C., Mickel, K.S., Moss, K.A., Kern H.F. and Scheele, G.A. (1987) EMBO J. 6, 43–48.
[12] Kozak, M. (1986) Cell 44, 283–292.
[13] Siflow, C.D., Chisholm, R.L., Conner, T.W. and Raneem, L.P.W. (1985) Mol. Cell Biol. 5, 2389–2398.
[14] Serrano, R. (1988) Biochim. Biophys. Acta 947, 1–28.
[15] Harper, J.F., Surowy, T.K. and Sussman, M.R. (1989) Proc. Natl. Acad. Sci. USA 86, 1234–1238.
[16] Ewing, N.N., Wimmers, L.E., Meyer, D.J., Chetelat, R.T. and Bennett, A.B. (1990) Plant Physiol. 94, 1874–1881.

[17] Perez, C., Michelet, B., Ferrant, V., Bogaerts, P. and Boutry, M. (1992) J. Biol. Chem. 267, 1204–1211.
[18] Addison, R. (1986) J. Biol. Chem. 261, 14896–14901.
[19] Serrano, R., Kielland-Brandt, M.C. and Fink, G.R. (1986) Nature 319, 689–693.
[20] Wolf, A.H., Slayman, C.W. and Gradmann, D. (1995) Plant Mol. Biol. 28, 657–666.
[21] Serrano, R. (1993) FEBS Lett. 325, 108–111.
[22] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Nucleic Acids Res. 22, 4673–4680.