Vol. 157, No. 1, 1988
November 30, 1988

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 203-209

cDNA CLONING AND SEQUENCE DETERMINATION OF PIG GASTRIC $(H^+ + K^+)$-ATPase

Masatomo Maeda, Jun Ishizaki and Masamitsu Futai

Department of Organic Chemistry and Biochemistry,
The Institute of Scientific and Industrial Research,
Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka 567, Japan

Received October 17, 1988

SUMMARY --- Complementary DNA to pig gastric mRNA encoding $(H^+ + K^+)$-ATPase was cloned, and its amino acid sequence was deduced from the nucleotide sequence. The enzyme contained 1034 amino acid residues (Mr. 114,285) including the initiation methionine. The sequence of pig $(H^+ + K^+)$-ATPase was highly homologous with that of the corresponding enzyme from rat, but had high degree of synonimous codon changes. Potential sites of phosphorylation by cAMP-dependent protein kinase and N-linked glycosylation sites were identified. The amino terminal region contained a lysine-rich sequence similar to that of the α subunit of $(Na^+ + K^+)$-ATPase, although a cluster of glycine residues was inserted into the sequence of the $(H^+ + K^+)$-ATPase. As the pig enzyme is advantageous for biochemical studies, the information of the primary structure is useful for further detailed studies.

INTRODUCTION --- The $(H^+ + K^+)$-ATPase of gastric parietal cell is located in its secretory canaliculi and catalyzes electroneutral exchange of internal H^+ and external K^+ coupled with ATP hydrolysis, maintaining a large pH difference of more than 5 between the cytoplasm and gastric lumen (1). Membrane vesicles enriched in this enzyme can be prepared from pig stomach easily and in large quantity (2,3), and the kinetic properties of the enzyme have been examined extensively (4-6). As the pig enzyme is advantageous for biochemical studies on coupled ion transport and its regulation, the information of the primary structure is prerequisite for further datailed studies. In this study, we cloned and sequenced cDNA coding for pig gastric $(H^+ + K^+)$-ATPase. From the deduced amino acid sequence, significance of the amino terminal region was discussed.

0006-291X/88 $1.50

MATERIALS AND METHODS

Poly(A)$^+$ mRNA was prepared from pig gastric mucosa. Double stranded cDNA was synthesized according to instructions in the manual of the cloning kit (Amersham) except that T_4 DNA polymerase treatment was omitted. The two 3'-ends of the double stranded cDNA were tailed with deoxycytidine, and PstI and SstI sites of pUC18 were tailed with deoxyguanosine using terminal deoxynucleotidyl transferase. They were then annealed and introduced into an E. coli strain C600 (7). A 36-base probe (^{5} TCCAAGTTTCCGGATCTCATCATAGACAAAGATGAG^3) corresponding to the sense strand of the 3'-terminal coding region of rat ($H^+ + K^+$)-ATPase (amino acid residues 1007-1018) (8) was 5'-end-labeled with T_4 polynucleotide kinase and used for colony hybridization (7). A hybridization-positive clone (pHK-I6) encoding the carboxyl terminal half of pig gastric ($H^+ + K^+$)-ATPase was obtained from among 10^4 clones. We found a sequence that was similar to that of the 36-base probe except for changes in two bases without change in the coded amino acid residues.

To obtain a cDNA clone carrying the amino-terminal region of the enzyme, cDNA was synthesized using a 30-base synthetic primer (^{5} CTGTTCGGCACTTGAGCACGGCATCAGGAA^3) corresponding to the sense strand of the 5'-terminal region (amino acid residues 609-619) of pHK-I6. The second library (10^4 clones) was probed with random-primed DNA using an equimolar mixture of 60- and 62-base synthetic oligonucleotides, (^{5} CAGTGCCTCATGTGGGTGGCTGCAGCCATCTGCCTCATTGCCTTTGC-GATTCAGGCCAGC^3 and ^{5} AATGAGTGCCAACGCCAGGTACAAATTGTCATCAGTGGTCAG-GTCTCCCTCGCTGGCCTGAA^3) corresponding to the antisense and sense strands of the amino terminal region of rat ($H^+ + K^+$)-ATPase (amino acid residues 110 - 129 and 126 - 146, respectively) (8) as templates, and a positive clone (pHK-M10) carrying the amino terminal half of the enzyme was obtained. No differences were found in the nucleotides in the overlapping regions of the two clones. A plasmid was prepared (7) and most of both strands of cDNA (94.9 %) were sequenced by the chain-termination method (9) following the sequence strategy shown in Fig.1.

Terminal deoxynucleotidyl transferase and random-primed DNA labeling kit were obtained from International Biotechnologies Inc. and Boehringer, respectively. Other enzymes were purchased from Takara Shuzo Co., Kyoto, Japan. $\gamma(^{32}P)$-ATP (5000 Ci/mmol) and $\alpha(^{32}P)$-dCTP (400 Ci/mmol and 3000 Ci/mmol) were products of the Radiochemical Centre, Amersham. All other chemicals used were of the highest grade commercially available.

RESULTS AND DISCUSSION

DNA sequencing of two hybridization-positive clones, pHK-I6 and pHK-M10 demonstrated that they contained the cDNAs coding carboxyl and amino terminal regions of pig ($H^+ + K^+$)-ATPase, respectively (Fig.1). Fig.2 shows the nucleotide sequence of cDNA encoding ($H^+ + K^+$)-ATPase determined using clones pHK-I6 and pHK-M10. The translational initiation site was assigned to the Met-1 codon, as the sequence around this codon (GCACCATGG) is consistent with the consensus sequence for eukaryotic initiator sites (10). A termination codon (TAG) was found next to the Tyr-1034 codon. The primary translation product consists of

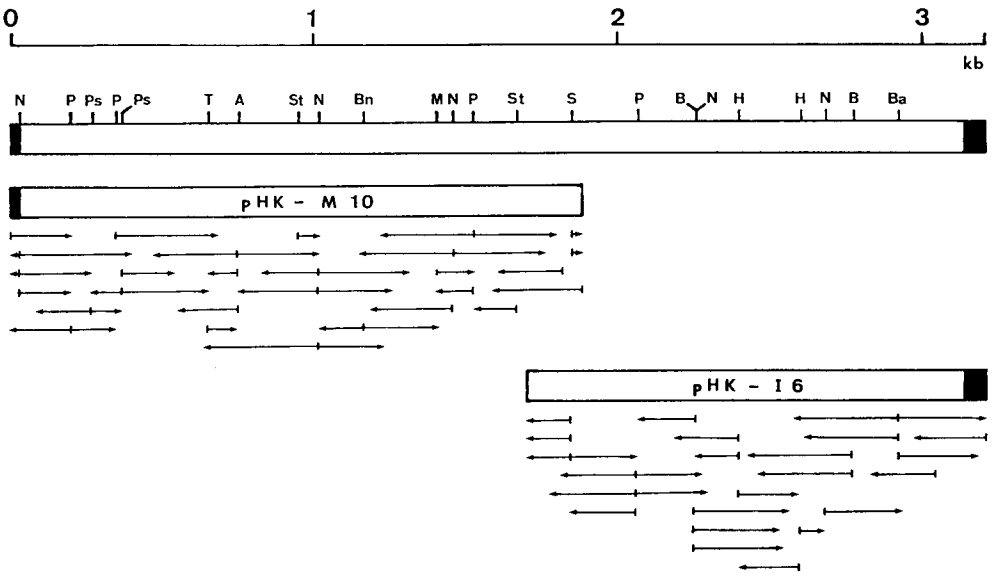

Fig.1. Restriction map and sequencing strategy for pig gastric $(H^+ + K^+)$-ATPase clones. The restriction map is a composite of those of the cDNAs whose sequences were determined. The restriction sites relevant to the sequencing strategy are shown: A, ApaLI; B, BalI; Ba, BamHI; Bn, BanII; H, HincII; M, MluI; N, NcoI; P, PvuII; Ps, PstI; S, SmaI; St, StuI; T, TaqI. Open areas represent the reading frames, and darkened areas represent untranslated regions. The directions and extents of sequencing are shown by arrows.

1034 amino acids and the molecular weight of the protein, including the initiator methionine, was calculated to be 114,285. We identified the sequences of peptides reported previously: the amino terminal sequence at positions 3 to 18 (11), the region around the phosphorylation site at positions 384 to 387 (CSDK) (12) and the region around the fluorescein isothiocyanate-binding site at positions 515 to 522 (LVMKGAPE) (13).

Alignment of the nucleotide sequences of pig and rat $(H^+ + K^+)$-ATPases indicated 89.1 % homology in the coding region, 60.0 % homology in the 5'-noncoding region and 44.3 % homology in the 3'-noncoding region. The amino acid sequence homology between the two enzymes was 97.6 % (identical residues) with 28.9 % synonymous codon changes for amino acid residues. The 4th amino acid residue (alanine) from the initiator methionine is not present in the rat enzyme. In addition, 24 other residues of the two enzymes are different. Of these, 16 residues are scored as favored substitutions (14). Substitutions are concentrated

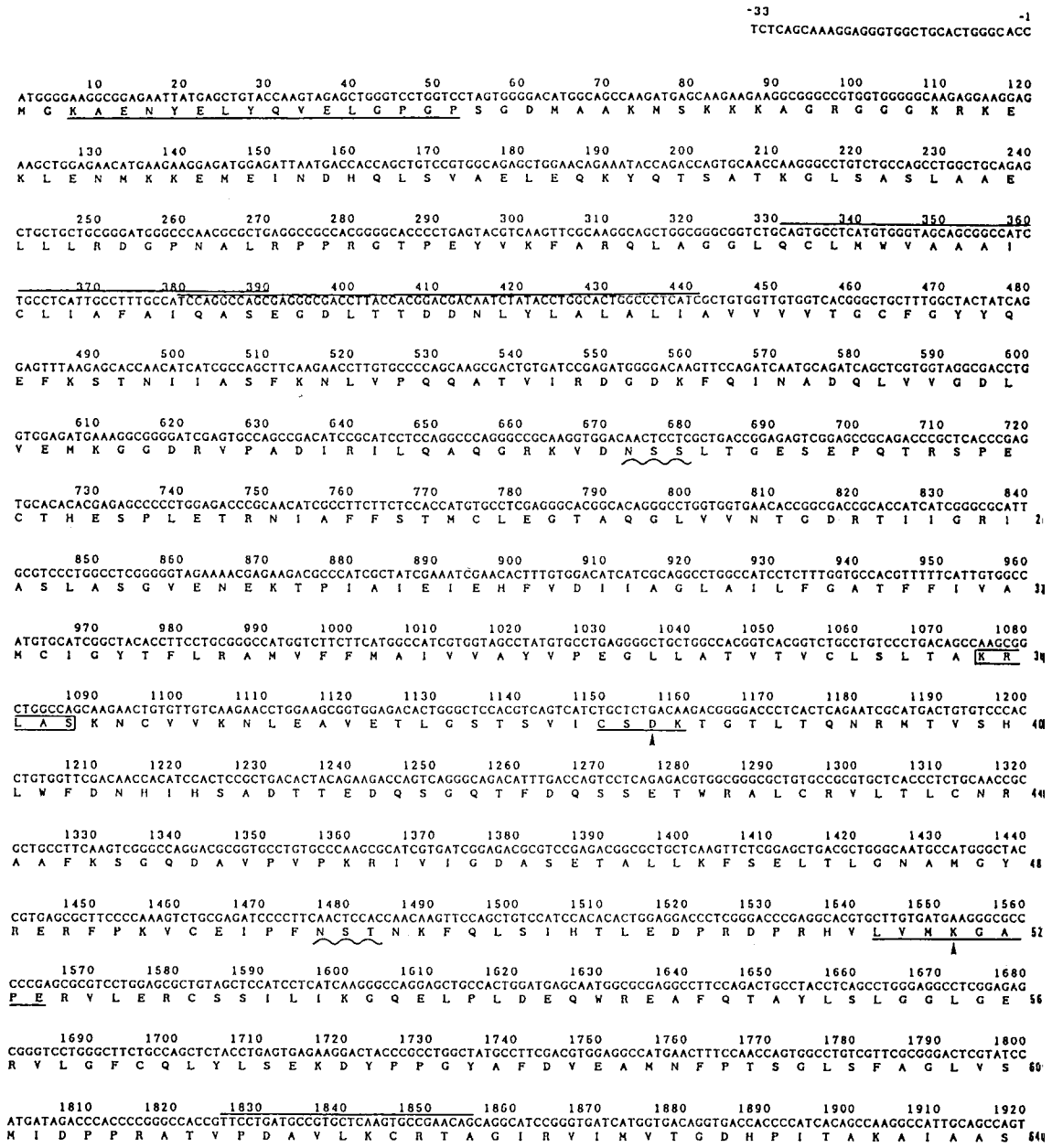

Fig.2. Nucleotide sequence of cDNA for pig gastric (H^+ + K^+)-ATPase and its deduced amino acid sequence. The nucleotide sequence was determined using clones pHK-M10 (amino acid residues 1-617) and pHK-16 (residues 557-1034). Nucleotides are numbered above the DNA sequence starting from the first letter of the initiation methionine, and numbers on the right of each line are those of amino acid residues starting from Met-1. Asp-386 (phosphorylated during the catalytic cycle) and Lys-518 (fluorescein isothiocyanate-binding site) are indicated by arrow heads. The consensus sequences of the N-glycosylation site and phosphorylation site by cAMP-dependent protein kinase are indicated by ~~~ and □, respectively. The amino acid sequences determined by protein chemical analysis are underlined. The lines above the nucleotide sequence correspond to the sequence of synthetic oligo-nucleotides described in the Materials and Methods.

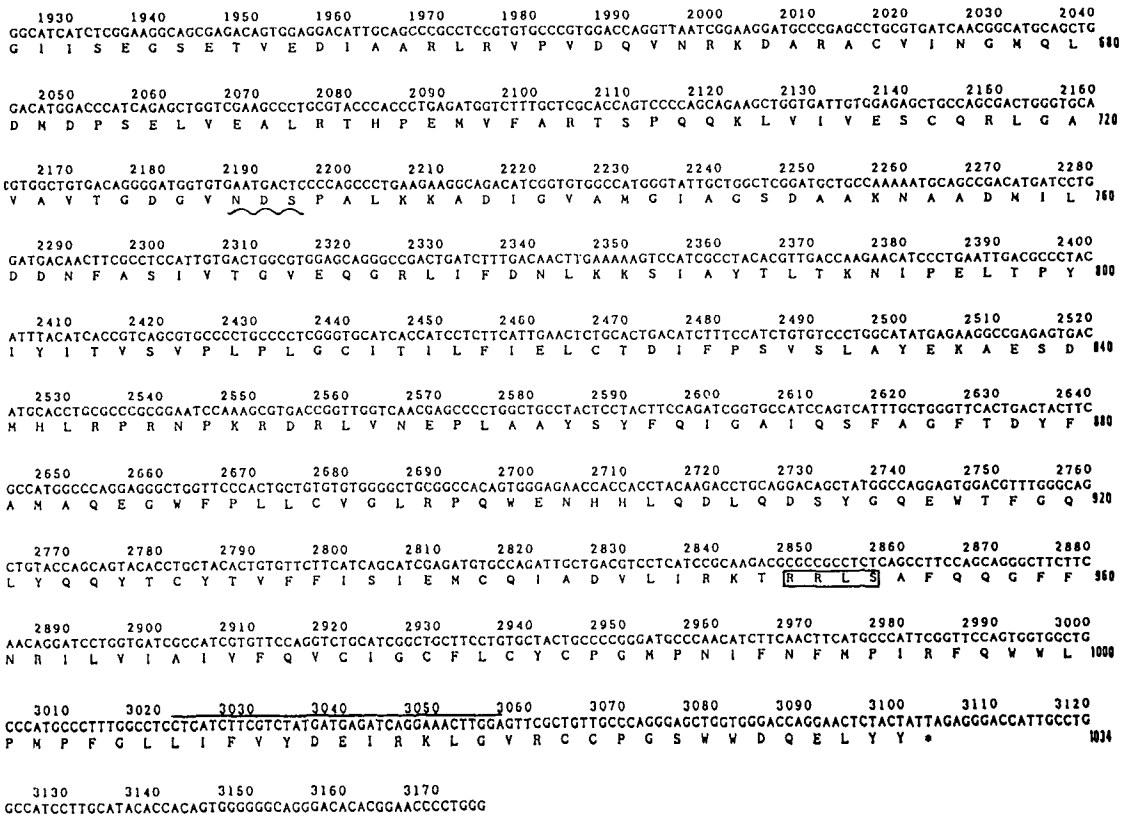

Fig. 2 - *Continued*.

in two regions (positions 1 to 75 and positions 570 to 600), in which 9 and 4 residues, respectively, are changed. In the α subunit of the closely related ($Na^+ + K^+$)-ATPase, 63.0 % of the amino acid residues are identical with those of pig ($H^+ + K^+$)-ATPase (15).

In the deduced amino acid sequence, consensus sequences for sites phosphorylated by cAMP-dependent protein kinase (KRXXS and RRXS) (16) were found at residues 359 to 363 (KRLAS) and residues 950 to 953 (RRLS), respectively. As histamine activates cAMP-dependent protein kinase, which is involved in the control of acid secretion from parietal cell (17), it will be interesting to know whether the enzyme activity is modulated by phosphorylation. Three Asn residues (Asn-225, 493, and 730) similar to known N-glycosylation sites (NX(S/T)) (18) were also found.

Gastric $(H^+ + K^+)$-ATPase is classified into cation transporting ATPases, such as $(Na^+ + K^+)$-, Ca^{2+}-, K^+- and H^+-ATPases (19). The primary structures of these enzymes are highly homologous, and conserved regions may be essential for common catalytic steps, such as ATP-binding, phosphorylation and dephosphorylation of conserved aspartic acid residue (Asp-386), and transmission of energy to cation transport. On the other hand, the cation selectivity may be determined by less homologous regions. $(H^+ + K^+)$-ATPases have a lysine-rich consensus sequence similar to that of the α subunits of $(Na^+ + K^+)$-ATPase in their amino terminal region (15, 20-25), $\mathbf{K_{1-2}X_{0-2}K_{1-3}X(GXGGG)(K/R)_{1-3}X(K/R)_{1-2}X_{1-4}Z_2XK_2E}$ where **Z** is **E**, **D** or **N**, although a cluster of glycine residues (**GXGGG**) is inserted in the middle of the sequence of $(H^+ + K^+)$-ATPase (Fig.3). Since this lysine-rich sequence is not found in other related cation transporting ATPases (19), it may have a common regulatory role(s) in cation binding and its occlusion by $(H^+ + K^+)$- and $(Na^+ + K^+)$-ATPases. The cluster of glycine residues found only in $(H^+ + K^+)$-ATPase may determine the enzyme specificity for H^+.

		1 10 20 30 40 50
H^+/K^+	Pig	MG KAENY ELYQVE LGPGPSGDMAAK MSKKKAGRGGGKRKEK LENMKKEME
	Rat	MG K ENY ELYSVE LGTGPGGDMAAK MSKKKAGGGGGKKKEK LENMKKEME
Na^+/K^+	Pig	MG KGVGR DKY EPAAVSEHGD K KKA KKER DMDELKKEVS
	Sheep	MG KGVGR DKY EPAAVSEHGD K KKA KKER DMDELKKEVS
	Chicken	MG KGAGR DKY EPTATSEHGT K KKKA KER DMDELKKEIS
	HeLa	MG KGVGR DKY EPAAVSEQGD KK G KKG KKDR DMDELKKEVS
	Rat α	MG KGVGR DKY EPAAVSEHGD KK S KKA KKER DMDELKKEVS
	α (+)	MG RGAGR E Y SPAATTAENGGG K KKQ KEK ELDELKKEVA
	α III	MGDK KDDK SSP KK S KA KERR DLDDLKKEVA
	Torpedo	MG KGAAS EKY QPAATSEN AK NS KKS KSK TTDLDELKKEVS
	Artemia	AK KKQ KKGK DLNELKKELD

Fig.3. Comparison of the N-terminal amino acid sequences of $(H^+ + K^+)$-ATPases from pig and rat and those of $(Na^+ + K^+)$-ATPase α-subunits from various species. The sequences of $(H^+ + K^+)$-ATPases shown are for pig (this study) and rat (8), and those of $(Na^+ + K^+)$-ATPase are for pig kidney α (15), sheep kidney α (20), chicken α (21), HeLa cells α (22), rat α, α(+), αIII (23), Torpedo californica electric organ α (24), and the brine shrimp Artemia salina (25). The residue numbers are based on the sequence for pig $(H^+ + K^+)$-ATPase. The first five amino acids of rat, sheep, and pig α and rat α(+) subunits are known to be removed by post-translational processing (15,20,23). Basic amino acid residues are shown in boldface. The line was drawn above the lysine-rich sequence.

ACKNOWLEDGMENTS: We are grateful to Prof. S. Nakanishi, Kyoto University, for valuable discussion. This research was supported in part by grants from the Ministry of Education, Science and Culture of Japan, Mitsubishi Foundation and the Foundation for Promotion of Pharmaceutical Science.

REFERENCES

1. Faller, L., Jackson, R., Malinowska, D., Mukidjam, E., Rabon, E., Saccomani, G., Sachs, G. and Smolka, A. (1982) Ann. N. Y. Acad. Sci. 402, 146-163.
2. Sachs, G., Chang, H. H., Rabon, E., Schackman, R., Lewin, M. and Saccomani, G. (1976) J. Biol. Chem. 251, 7690-7698.
3. Maeda, M., Tagaya, M. and Futai, M. (1988) J. Biol. Chem. 263, 3652-3656.
4. Wallmark, B., Stewart, H. B., Rabon, E., Saccomani, G. and Sachs, G. (1980) J. Biol. Chem. 255, 5313-5319.
5. Stewart, B., Wallmark, B. and Sachs, G. (1981) J. Biol. Chem. 256, 2682-2690.
6. Rabon, E. C., McFall, T. L. and Sachs, G. (1982) J. Biol. Chem. 257, 6296-6299.
7. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.).
8. Shull, G. E. and Lingrel, J. B. (1986) J. Biol. Chem. 261, 16788-16791.
9. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. and Roe, B. A. (1980) J. Mol. Biol. 143, 161-178.
10. Kozak, M. (1984) Nucleic Acids Res. 12, 857-872.
11. Lane, L. K., Kirley, T. L. and Ball, W. J. Jr. (1986) Biochem. Biophys. Res. Commun. 138, 185-192.
12. Walderhaug, M. O., Post, R. L., Saccomani, G, Leonard, R. T. and Briskin, D. P. (1985) J. Biol. Chem. 260, 3852-3859.
13. Farley, R. A. and Faller, L. D. (1985) J. Biol. Chem. 260, 3899-3901.
14. Dayhoff, M. O., Schwartz, R. M. and Orcutt, B. C. (1978) Atlas of Protein Sequence and Structure (Dayhoff, M. O. ed.) Vol.5, Suppl.3, pp.345-352, (National Biomedical Research Foundation, Silver Spring, Md.).
15. Ovchinnikov, Y. A., Modyanov, N. N., Broude, N. E., Petrukhin, K. E., Grishin, A. V., Arzamazova, N. M., Aldanova, N. A., Monastyrskaya, G. S. and Sverdlov, E. D. (1986) FEBS Lett. 201, 237-245.
16. Krebs, E. G. and Beavo, J. A. (1979) Ann. Rev. Biochem. 48, 923-959.
17. Chew, C. S. (1985) J. Biol. Chem. 260, 7540-7550.
18. Hubbard, S. C. and Ivatt, R. J. (1981) Ann. Rev. Biochem. 50, 555-583.
19. Serrano, R. (1988) Biochim. Biophys. Acta 947, 1-28.
20. Shull, G. E., Schwartz, A. and Lingrel, J. B. (1985) Nature 316, 691-695.
21. Takeyasu, K., Tamkun, M. M., Renaud, K. J. and Fambrough, D. M. (1988) J. Biol. Chem. 263, 4347-4354.
22. Kawakami, K., Ohta, T., Nojima, H. and Nagano, K. (1986) J. Biochem. (Tokyo) 100, 389-397.
23. Shull, G. E., Greeb, J. and Lingrel, J. B. (1986) Biochemistry 25, 8125-8132.
24. Kawakami, K., Noguchi, S., Noda, M., Takahashi, H., Ohta, T., Kawamura, M., Nojima, H., Nagano, K., Hirose, T., Inayama, S., Hayashida, H., Miyata, T. and Numa, S. (1985) Nature, 316, 733-736.
25. Morohashi, M. and Kawamura, M. (1984) J. Biol. Chem. 259, 14928-14934.