SEQUENCE CONSERVATION OF THE CATALYTIC REGIONS OF AMYLOLYTIC ENZYMES IN MAIZE BRANCHING ENZYME-I

Tadashi Baba*, Koji Kimura, Kouichi Mizuno, Hirotoshi Etoh, Yoshiki Ishida, Osamu Shida**, and Yuji Arai

Institute of Applied Biochemistry, University of Tsukuba, Ibaraki 305, Japan

Received October 14, 1991

We have identified cDNA clones encoding branching enzyme-I (BE-I) from a maize kernel cDNA library. The combined nucleotide sequence of the cDNAs indicates that maize BE-I is initially synthesized as a precursor protein with a putative 64-residue transit peptide at the amino terminus, and that the mature enzyme contains 759 amino acid residues with a calculated molecular mass of 86,236 Da. The four regions, which constitute the catalytic site of amylolytic enzymes, are conserved in the sequences of BE-I and bacterial branching enzymes. This result demonstrates that branching enzyme belongs to a family of the amylolytic enzymes. The BE-I gene is highly expressed in the early stages of kernel development, and the level of the message concentration decreases slowly as kernel maturation proceeds. $_{0.1991~\text{Academic}}$ Press, Inc.

In the storage organs of higher plants, starch is synthesized and accumulated in the organelles, amyloplast. Starch is composed of two chemically different molecules, amylose and amylopectin; amylose is an essentially linear polymer with $\alpha-1,4$ -glucosidic bonds, whereas amylopectin is a highly branched molecule containing $\alpha-1,4$ - and $\alpha-1,6$ -linkages. It has long been considered that biosynthesis of starch in the amyloplast is catalyzed by starch synthase, branching enzyme, and possibly phosphorylase (1). However, the synthetic mechanisms of amylose and amylopectin are poorly understood. Therefore, to elucidate the mechanisms at the protein and gene levels, it is essential to isolate and characterize the genes for starch-synthesizing enzymes.

Plant branching enzyme acts on already synthesized and/or elongating amylose chains to form $\alpha-1,6-1$ inked branch points. Thus, this enzyme is a key enzyme for amylopectin biosynthesis (5-8). In maize kernels, two types

 $^{{}^*\}mathrm{To}$ whom correspondence should be addressed.

Present address: Research Laboratory, Higeta Shoyu Co. Ltd., Choshi, Chiba 288, Japan.

of branching enzyme, BE-I and BE-II, have been identified (5-7). Recently, Bhattacharyya et al. (9) reported the cloning of the gene encoding an isoform of branching enzyme in pea. Although the wrinkled-seed character of pea is found to be caused by a transposon-like insertion in the branching enzyme gene (9), the sequence data for the coding region of the gene has not been available yet.

We now describe the amino acid sequence of maize BE-I deduced from the cDNA sequence, and the gene expression during kernel development. This is the first report on the primary structure of branching enzyme in plant.

MATERIALS AND METHODS

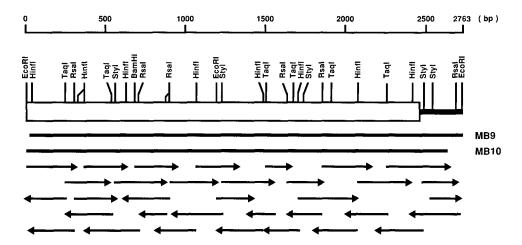
<u>Materials.</u> Maize ($\underline{\text{Zea}}$ $\underline{\text{mays}}$ L.) inbred 0h43 was field-grown in 1988. All plants were self- or sib-pollinated, and the ears were harvested 10 to 28 days after pollination, immediately frozen in liquid nitrogen, and stored at -80°C. BE-I was purified according to the previously described procedure (8). The purified enzyme migrated as a protein with a molecular mass of 85 kDa on SDS-PAGE.

Preparation of antibodies. The purified BE-I (50 μ g) was mixed with an equal volume of Freund's complete adjuvant, and injected intradermally into New Zealand White rabbits. The injection was repeated twice at intervals of two weeks using Freund's incomplete adjuvant. Antisera were collected two weeks after the final injection. The immunoglobulin fraction was prepared by precipitation with 50% saturated ammonium sulfate, and subjected to immunoaffinity chromatography on a BE-I-conjugated Sepharose 4B column. The antibodies retained on the column were eluted with 0.2 M glycine/HCl, pH 2.2, immediately neutralized with 0.4 M potassium phosphate, pH 8.0, and then dialyzed against phosphate-buffered saline.

<u>Preparation of RNA.</u> Kernels (20 g) were ground with a mortar and pestle in liquid nitrogen, homogenized with a Polytron in 200 ml of 8 M urea containing 3 M LiCl, and allowed to stand overnight at 4°C. After centrifugation at $16,000 \times g$ for 20 min, the precipitate was dissolved in 50 ml of 10 mM Tris/HCl, pH 7.6, containing 0.5% SDS, and treated with phenol/chloroform solution. Total RNA was recovered by ethanol precipitation. Polyadenylated RNA was selected by oligo(dT)-cellulose (Type 7, Pharmacia LKB Biotechnology) column chromatography.

Construction and screening of a maize kernel cDNA library. Polyadenylated RNA from 14-day-old kernels was denatured with 10 mM methylmercury hydroxide for 5 min at room temperature. Double-stranded cDNA was synthesized from the denatured RNA using a commercial kit of Pharmacia LKB Biotechnology. The cDNA was ligated into the EcoRI site of λ gtll. After in vitro packaging, 130,000 plaques were screened by affinity-purified anti-BE-I antibodies, as described previously (10). Positive clones were plaque-purified, and the cDNA inserts were subcloned into pUC19 or M13mp19 vector. Nucleotide sequencing was carried out by the dideoxy chain termination method (11), using a Sequenase Version 2.0 (U.S.B. Corp.) or an AmpliTaq sequencing kit (Takara).

Northern blot analysis. Total RNA was denatured with glyoxal, separated by electrophoresis on 1.2% agarose gels, and transferred onto GeneScreen Plus nylon membranes. The blot was probed by the 32 P-labeled cDNA fragment



<u>Fig. 1.</u> Restriction map and sequencing strategy for cDNA inserts encoding maize branching enzyme-I (BE-I). The scale at the top designates nucleotide position in base pair (bp) from the 5'-end of the MB10 insert. The open box indicates the protein coding region including a putative transit peptide. The arrows indicate the direction and extent of nucleotide sequence determined from each site.

of MB9 (Fig. 1). Stringencies employed for hybridization and washing were the same as those described previously (12).

<u>Protein analysis.</u> Purified BE-I (0.1 mg of protein) was lyophilized, dissolved in 70% formic acid (0.2 ml) containing 0.1 mg CNBr, and incubated at room temperature for 2 days in the dark. The mixture was applied on a LiChrosphere RP-8 reverse phase column (Cica-Merck) previously equilibrated with 0.1% trifluoloacetic acid. Peptides were eluted with a linear gradient (0-60%) of acetonitrile in the same solution, collected manually, and lyophilized. The original BE-I and the two CNBr fragments were sequenced by automated Edman degradation using an Applied Biosystems Model 477A protein sequencer equipped with a Model 120A on-line PTH analyzer.

RESULTS AND DISCUSSION

When a maize kernel cDNA library was screened by affinity-purified anti-BE-I antibodies, four positive clones, termed MB9, MB10, MB18, and MB19, were obtained. Restriction mapping and partial nucleotide sequencing of these cDNA inserts indicated that they were all related to one another. Of the four clones, the cDNA inserts of MB9 and MB10 were completely sequenced (Fig. 1). The combined nucleotide sequence of the overlapping cDNAs was 2,763 base pairs (bp) in length and contained an open reading frame of 2,469 nucleotides (Fig. 2). The deduced amino acid sequence at residues 1-6 matched the amino-terminal sequence of BE-I (Ala-Thr-Val-Gln-Glu-Asp-) determined by protein analysis. The sequences of two CNBr fragments from BE-I were also found at residues 10-29 and 534-546 (underlines in Fig. 2). These

results clearly demonstrate the identities of these inserts as the BE-I cDNA clone. Since the nucleotide sequence lacked an ATG initiation codon, rescreenings of both the maize kernel cDNA and genomic libraries were carried out using the cDNA fragment of MB9 as a probe. Positive clones were practically isolated, but failed to complete the open reading frame. At any rate, the mature BE-I contains 759 amino acid residues with a calculated molecular mass of 86,236 Da. This value agrees well with that obtained experimentally by SDS-PAGE (8).

The amino acid sequence of BE-I at residues -63 to -1 preceding the amino-terminal amino acid (Ala^1) is significantly different from the signal sequences of secretory proteins in length and amino acid composition (Fig. 2). It is widely known that precursors of nuclear-encoded chloroplastic proteins have transit peptides at their amino termini (13). The function of the transit peptide is believed to be the transport and/or import of the precursors into the organella. The 63-residue sequence as the amino-terminal extension of BE-I is rich in basic (21% as Arg, Lys, and His) and hydroxylic (19% as Ser and Thr) amino acids, and contains only an acidic amino acid (Asp^{-40} , see Fig. 2). This feature is consistent with the characteristic of the transit peptide of potato phosphorylase precursor which is routed into the amyloplast (14). Thus, the amino-terminal extension of BE-I most likely acts as a transit peptide for the amyloplast.

We have recently identified a cDNA clone, termed RB13, encoding rice BE-I (the details will be reported elsewhere). The insert of RB13 contains an almost full-length cDNA which includes an initiator ATG codon. Comparison of the amino-terminal extensions between maize and rice BE-I shows an extensive sequence similarity (55% identity); especially, the 7 amino acids of the amino-terminal 9-residue sequence (at residues -63 to -55 in maize BE-I, see Fig. 2) are exactly identical between maize and rice. This suggests that the maize sequence lacks only the initiator methionine from the full protein coding region. In fact, Northern blot analysis yields a 2.8-kb signal as the transcript size corresponding to the overlapping cDNA inserts of MB9 and MB10, as described below. Therefore, it is reasonable to consider that maize BE-I is initially synthesized as a 823-residue precursor protein, including a 64-residue transit peptide.

Fig. 2. Nucleotide sequence of maize branching enzyme-I (BE-I) cDNA and its deduced amino acid sequence. The deduced amino acid sequence is shown by the standard one-letter code below the nucleotide sequence numbered in the 5' to 3' direction. The amino-terminal sequences of BE-I and its two CNBr fragments determined by protein analysis are underlined. The amino acids are numbered from the amino terminus of the mature enzyme, and the residues on the amino-terminal side of residue I (Ala) are indicated by negative numbers. A putative polyadenylation signal is broken-underlined.

в ств тес стс 10 -61 GTG TOG COC TICT TOC TOG COG ACT COG CIT COG COG COG COG COG COC TICT COC TOG CAT GAT COG GCG GCA COG COG GGG ATC GCG GGT 100 S P S S S P T P L P P P R R S R S H A D R A A P P GEC GEC AAT GTG CEC CTG AGT GTG TTG TCT GTC CAG TGC AAG GCT CEC CEG TCA GEG GTG CEG AAG GTC AAG AGC AAA TTC GCC ACT GCA 190 V O C K A R R S G V R K V 280 30 TTC AMG GAC CAT TTC AGG TAC CGG ATG AMA AGA TTC CTA GAG CAG AMA GGA TCA ATT GAA GAA AAT GAG GGA AGT CTT GAA TCT TTT TCT F K D H F R Y R M K R F L E Q K G S I E E N E G S L E S F S370 60 AMA GOC TAT TIG AMA TIT GOG ATT MAT ACA MAT GAG GAT GOA ACT GTA TAT COT GAA TOG GAC COT GCT GCC CAG GAG GCA GAG CTT ATT K G Y L K F G I N T N E D G T V Y R E W A P A A Q E A E L I460 GET GAC TIC AAT GAC TIGG AAT GET ECA AAC CAT AAG ATG GAG AAG GAT AAA TIT GET GIT TIGG TICG ATC AAA ATT GAC CAT GTC AAA GEG N D W N G A N H K M E K D K F G V W S AMA OCT GOC ATC OCT CAC AAT TOC AAG GTT AMA TIT COC TIT CTA CAT GGT GGA GTA TGG GTT GAT CGT ATT CCA GCA TTG ATT CGT TAT 640 150 Gog act sit gat goc ict ama tit gea got goc tat gat get sit cat igg gat oct oct sct ict gan agg tac aca tit amg cat cct A T V D A S K F G A P Y D G V H W D P P A S E R Y T F K H P730 180 dgg oct tca aag oct gct gct oca ogt atc tat gaa goc cat gta agt agt agt gat gaa aag oca gca gta agc aca tat agg gaa tit 820 EAHVGMSGEKPA 210 910 GCA GAC AAT GTG TTG OCA OGC ATA OCA GCA AAT AAC TAC AAC ACA GTT CAG TTG ATG GCA GTT ATG GAG CAT TOG TAC TAT GCT TCT TTC
A D N V L P R I R A N N Y N T V Q L M A V M E H S Y Y A S F 240 Geg tac cat etg aca aat tite tit egg ett age aga tea eeg aca ca ca eag gac etc aaa tat ett ett git gat aag eca cac agi tig ${\tt G}$ ${\tt Y}$ ${\tt H}$ ${\tt V}$ ${\tt T}$ ${\tt N}$ ${\tt F}$ ${\tt F}$ ${\tt A}$ ${\tt V}$ ${\tt S}$ ${\tt S}$ ${\tt R}$ ${\tt S}$ ${\tt G}$ ${\tt T}$ ${\tt P}$ ${\tt E}$ ${\tt D}$ ${\tt L}$ ${\tt K}$ ${\tt Y}$ ${\tt L}$ ${\tt V}$ ${\tt D}$ ${\tt K}$ ${\tt A}$ ${\tt H}$ ${\tt S}$ ${\tt L}$ 1000 270 GET THE CEA STT CHE ATE GAT GIT GEC CAT AGC CAT GCA AGT AAT AAT GTC ACA GAT GET TTA AAT GGC TAT GAT GIT GEA CAA AGC ACC 1090 H S H A S N N V T D G L N สกก caa gag toc tat titt cat gog gga gat aga ggt tat cat aaa cit tigg gat agt cog cig titc aac tat gct aac tigg gag gita tita agg 1180 HAGDRGYHKLWDSRLF TIT CIT ICT ACC CIG AGA TAT TEG TITG GAT GAA TIC ATG TIT GAT GEC TITC GEA TITT GAT GEA GIT ACA TCA ATG CIG TAT CAT CAC 1270 D 360 1360 1450 420 AAC CAT TTA ATG CAC AAA CTC TTG CCA GAA GCA ACT GTT GTT GCT GAA GAT GTT TCA GGC ATG COG GTC CTT TGC CGG CCA GTT GAT GAT NHLMHKLLPEATVVAEDVSGMP 1540 450 atg get gan ata gog cat act tig act aac agg aga tat act gan ama tec atc gca tat gct gag agc cat gat cag tct att git gog M G E I A H T L T N R R Y T E K C I A Y A E S H D Q S I V G 1630 480 gac ana act att gca titt ctc ctg atig gac aag gaa atig tac act ggc atig tica gac titg cag oct gct tica oct aca att gat oga ggg 510 ATT GCA CITC CAA AAG ATG ATT CAC TTC ATC ACA ATG GCC CIT GGA GGT GAT GGC TAC TTG AAT TTT ATG GGA AAT GAG TTT GCT CAC CCA 1810 GAA TGG ATT GAC TIT CO'A AGA GAA GGG AGC AAC TGG AGC TAT GAT AAA TGC AGA CGA CAG TGG AGC CIT GTG GAC ACT GAT CAC TTG CGG E W I D F P R E G N N W S Y D K C R R O W S L V D T D H L R 1900 DFPREGNN 570 1990 TAC ANG TAC ATG AAT GOG TIT GAC CAA GOG ATG AAT GOG CTC GAT GAG AGA TIT TOC TTC CTT TOG TOG TCA AAG CAG ATC GTC AGC GAC $\,^{\prime}$ $\,^{\prime$ 2080 ATG AAC GAT GAG GAA AAG GIT ATT GIC TIT GAA OGT GGA GAT ITA GIT TIT GIT TIC AAT TIC CAT OOC AAG AAA ACT TAC GAG GGC TAC 630 V F E R G D L V F AMA GTIG GEA TEC GAT TTIG COT GEG AMA TAC AGA GTA GCC CTIG GAC TCT GAT GCT CTIG GTC TTC GGT GEA CAT GEA AGA GTT GGC CAC GAC K V G C D L P G K Y R V A L D S D A L V F G G H G R V G H D2170 660 2260 690 2350 COC COC ACC TET STG GCT TAT TAC CET STA GAC GAA GCA GGG GCT GGA CGA CET CTT CAC GCG AAA GCA GGA AAG AGG ACA GGA AAG ACG TCT CCA CVAYYRVDEAGAGRRLHAKAE 720 2440 750 AGOCACGAGT OCTTOGTIGAG GACTEGACTG GCTGCCGGCG CCCTGTTAGT AGTCCTGCTC TACTGGACTA 2540 OGG CAG CCA TOC GAT CAA GAT ACC AAA TGA 759 OPSDODTK GOGGOGGET GOGGOCCTTGG AACGGTCCTT TCCTGTAGCT TGCAGGOGGC TGGTGTCTCA TCACCGAGCA GGCAGGCACT GCTTGTATAG CTTTTCTAGA ATAATAATCA 2650 GGGATEGATG CATGETGTGT ATTGGCTATC TGGCTAGAGG TGCATGTGCC CAGTTTGTAT GTACAGGAGC AGTTCCCGTC CAGAATAAAA AAAACTTGT TGGGGGGGTTT 2760 2763

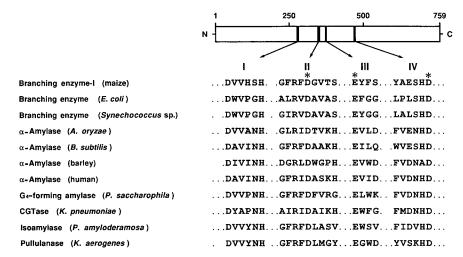


Fig. 3. Comparison of the deduced amino acid sequence of maize branching enzyme—I (BE-I) with those of bacterial branching enzymes and typical amylolytic enzymes. The sequences of maize BE-I at residues 277-282, 347-355, 373-376, and 470-475 (see Fig. 2) are compared with those of the four conserved regions (Regions I, II, III, and IV) of typical amylolytic enzymes, as described by Nakajima et al. (19): α -amylases from Aspergillus oryzae (20), Bacillus subtilis (21), barley (22), and human (23), Pseudomonas saccharophila maltotetraohydrolase (G4-forming amylase, ref. 24), Klebsiella pneumoniae cyclodextrin-glycosyltransferase (CGTase, ref. 25), Pseudomonas amyloderamosa isoamylase (26), and Klebsiella aerogenes pullulanase (27). The sequences of branching enzymes from Escherichia coli (15) and Synechococcus sp. (16) are also aligned. The possible catalytic residues for cleavage of α -1,4-linkages (17, 18) are indicated by asterisks.

The genes for two bacterial branching enzymes have been cloned and sequenced (15, 16). A significant similarity was found in the amino acid sequences between these two proteins (46% identity, see ref. 16). When the sequence of maize BE-I was compared with those of the bacterial enzymes, the similarity was low (approximately 23% identities for both bacterial enzymes). Also, the sequence similarity of BE-I with various amylolytic enzymes was quite limited. However, the four regions (Regions I, II, III, and IV), which constitute the catalytic site of the amylolytic enzymes, including isoamylase and pullulanase that split $\alpha-1,6$ -glucosidic bonds in the branched glucans, are conserved in the sequence of maize BE-I as well as of bacterial branching enzymes (Fig. 3). Interestingly, plant and bacterial branching enzymes all conserve the possible catalytic residues that are required for cleavage of $\alpha-1,4$ -glucosidic bond, including aspartate, glutamate, and aspartate (17, 18) in Regions II, III, and IV, respectively (residues 351, 373, and 475 in the BE-I sequence, see Figs. 2 and 3). These data demonstrate that branching enzyme belongs to a family of the amylolytic enzymes.

The developmental pattern of BE-I gene expression was examined using total RNAs from kernels at 10 to 28 days after pollination (Fig. 4). The

10 14 18 22 28 DAP

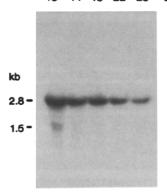


Fig. 4. Expression of the maize branching enzyme-I (BE-I) gene during kernel development. Total RNA was isolated from kernels at 10 to 28 days after pollination (DAP). A Northern blot was prepared, using 20 μ g each of these RNAs, and probed by the 32 P-labeled cDNA insert of MB9 (nucleotides 19-2763, see Figs. 1 and 2), as described in MATERIALS AND METHODS.

2.8-kb BE-I message was found abundantly at the 10th day of kernel The level of the message concentration decreased slowly as development. kernel maturation proceeded. A transcript with an approximate size of 1.5 kb, which was hybridized with the BE-I cDNA, was also detectable (Fig. 4). The 1.5-kb mRNA concentration was more abundant in the early stages than in the later stages of kernel development. However, it is not clear at present what is encoded by the 1.5-kb message.

In this paper, we describe the sequence conservation of the catalytic site of amylolytic enzymes in maize BE-I. This fact will be an important clue for studies of the structure and function of the enzyme. In addition, experiments are underway to define the roles served by the transit peptide of the BE-I precursor.

REFERENCES

- 1. Preiss, J., and Levi, C. (1980) In the Biochemistry of Plants (Preiss, J., ed), pp. 371-423. Academic Press, New York.
- Drummond, G. S., Smith, E. E., and Whelan, W. J. (1972) Eur. J. Biochem. 26, 168–176.
- Borovsky, D., Smith, E. E., and Whelan, W. J. (1975) Eur. J. Biochem. 59, 615–625.
- 4. Borovsky, D., Smith, E. E., Whelan, W. J., French, D., and Kikumoto,
- S. (1979) Arch. Biochem. Biophys. 198, 627-631.
 Boyer, C. D., and Preiss, J. (1978) Carbohydr. Res. 61, 321-334.
 Boyer, C. D., and Preiss, J. (1978) Biochem. Biophys. Res. Commun. 80, 169-175.
- 7. Singh, B. K., and Preiss, J. (1985) Plant Physiol. 79, 34-40.
- 8. Baba, T., Arai, Y., Ono, T., Munakata, A., Yamaguchi, H., and Itoh, T. (1982) Carbohydr. Res. 107, 215-230.

- Bhattacharyya, M. K., Smith, A. M., Ellis, T. H. N., Hedley, C., and
- Martin, C. (1990) Cell 60, 115-122.

 Huynh, T. V., Young, R. A., and Davis, R. W. (1985) In DNA Cloning (Glover, D. M., ed) vol. 1, pp. 49-78, IRL Press, Oxford.
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463-5467.
- 12. Baba, T., Kashiwabara, S., Watanabe, K., Itoh, H., Michikawa, Y., Kimura, K., Takada, M., Fukamizu, A., and Arai, Y. (1989) J. Biol. Chem. 264, 11920-11927.
- 13. Schmidt, G. W., and Mishkind, M. L. (1986) Ann. Rev. Biochem. 55, 879-912.
- Nakano, K., Mori, H., and Fukui, T. (1989) J. Biochem. 106, 691-695.
- Baecker, P. A., Greenberg, E., and Preiss, J. (1986) J. Biol. Chem. 261, 8738-8743.
- 16. Kiel, J. A. K. W., Boels, J. M., Beldman, G., and Venema, G. (1990) Gene 89, 77-84.
- Matsuura, Y., Kusunoki, M., Harada, W., and Kakudo, M. (1984) J. Biochem. 95, 697-702.
- Buisson, G., Duee, E., Haser, R., and Payan, F. (1987) EMBO J. 6, 18. 3909-3916.
- 19. Nakajima, R., Imanaka, T., and Aiba, S. (1986) Appl. Microbiol. Biotech. 23, 355-360.
- Toda, H., Kondo, K., and Narita, K. (1982) Proc. Japan Acad. 58, 208-212. 20.
- Yang, M., Galizzi, A., and Henner, D. (1983) Nucl. Acids Res. 11, 237-249.
- 22. Rogers, J. C., and Milliman, C. (1983) J. Biol. Chem. 258, 8169-8174.
- 23. Nakamura, Y., Ogawa, M., Nishide, T., Emi, M., Kosaki, G., Himeno, S., and Matsubara, K. (1984) Gene 28, 263-270.
- Zhou, J., Baba, T., Takano, T., Kobayashi, S., and Arai, Y. (1989) FEBS 24. Lett. 255, 37-41.
- Binder, F., Huber, O., and Bock, A. (1986) Gene 47, 269-277.
- Amemura, A., Chakraborty, R., Fujita, M., Noumi, T., and Futai, M. (1988) J. Biol. Chem. 263, 9271-9275.
- 27. Katsuragi, N., Takizawa, N., and Murooka, Y. (1987) J. Bacteriol. 169, 2301-2306.