FEBS Letters 354 (1994) 123–127

FEBS 14718

Characterisation of a complementary DNA encoding a novel plant enzyme with sucrolytic activity

Gordon C. Machray[a,*], L. Burch[c], P.E. Hedley[a], H.V. Davies[b], R. Waugh[a]

[a]*Cell and Molecular Genetics Department, Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, Scotland, UK*
[b]*Cellular and Evironmental Physiology Department, Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, Scotland, UK*
[c]*Department of Biochemistry, University of Dundee, Dundee DD1 4HN, Scotland, UK*

Received 17 August 1994; revised version received 28 September 1994

Abstract The cloning of a 1332 bp cDNA from a potato (*Solanum tuberosum*L.) cv. Cara leaf cDNA expression library, using an antibody raised against a purified tuber protein preparation with sucrolytic activity, is described. The corresponding gene in potato is of low copy number, is expressed in a variety of tissues, and encodes a protein which includes several domains with similarity to database sequences, including ferredoxin from *Clostridium pasteurianum*. Expression of the cDNA in *E. coli* yields a fusion protein with sucrolytic activity.

Key words: Sucrose cleavage; Potato; *Solanum tuberosum*; Ferredoxin

1. Introduction

The disaccharide sugar sucrose, a major early product of photosynthesis, is pivotal to a variety of fundamental processes in plant metabolism (see [1] for a recent review). It can form an important store of carbohydrate, both in the short term and for longer periods, within the plant cell. It can also act as a substrate or precursor for the synthesis of storage polysaccharides, such as starch and fructans. Sucrose, in the majority of plants, is the medium for transport of photosynthate from source to sink tissue via the phloem system with the potential for both symplastic and apoplastic pathways. These wide-ranging and important functions are effected by only a limited number of enzymes known to catalyse metabolic conversions of sucrose. Sucrose cleavage in plants has previously been assigned to two different enzymes – invertase (β-fructofuranosidase, E.C. 3.2.1.26) and sucrose synthase (UDP-glucose:fructose glycosyl transferase, E.C. 2.4.1.13) – each of which has several isoforms with differing subcellular locations [2]. Fructosyl transferases, such as sucrose-sucrose fructosyl transferase, metabolise sucrose in the biosynthetic reactions of fructan formation [3]. Knowledge of the distribution and activity of these enzymes within the plant is essential to an understanding of carbohydrate metabolism. We have previously reported the cloning of invertase cDNAs and purification of invertase proteins from potato [4–6]. This work has led to the description here of a further novel plant enzyme with the ability to cleave sucrose.

2. Materials and methods

2.1. Plant material, library construction and screening

A cDNA library (in λ-ZAP) from tuber sprouts of potato cv. Dianella was prepared commercially and screened using a polyclonal antibody raised against a protein preparation with sucrolytic activity purified from potato tubers [6] using the *pico*BLUE™ immunoscreening protocol (Stratagene Ltd, La Jolla, CA). After plaque purification, positive clones were subjected to *in vivo* excision. The largest insert (1,111 bp) from the characterised plasmids was then used as a hybridisation probe to screen the above cDNA library and a cDNA library prepared in λ-gt11 from leaf mRNA of potato cv. Cara as described previously [4]. The largest insert (1,332 bp) obtained in a positive clone from the leaf cDNA library, was subcloned into pUC19 to yield plasmid pEF141 and sequenced on both strands by the dideoxy method [7] using Sequenase v.2.0 (US Biochemical, Cleveland, OH) with progressive oligonucleotide primer design.

2.2. Southern and gene expression analyses

Southern blotting and hybridisation using Hybond N^+ (Amersham International, Little Chalfont, England) were performed as described in the manufacturer's recommendations using 20 μg DNA from potato cv. Saturna leaf per lane, digested with the indicated restriction enzyme and probed using the 1,332 bp insert from pEF141. Expression analysis was conducted by RT-PCR as described [8].

2.3. Preparation of construct expressing fusion protein and protein purification and assay

The indicated cDNA sequence was subcloned into pGEX-2T (Pharmacia LKB Biotechnology, Uppsala, Sweden) from which a fusion protein was expressed and purified using glutathione sepharose 4B as recommended by the manufacturer. The iron and inorganic sulfide content of protein was assayed as for clostridial ferredoxins [9]. SDS-polyacrylamide gel electrophoresis and activities of the purified protein using sugar substrates were assayed as detailed [6].

3. Results and discussion

3.1. Nucleotide and derived amino acid sequence

The entire sequence of the 1,332 bp cDNA cloned in pEF141 is presented in Fig. 1. The identity between this and the partial cDNA first isolated was confirmed by limited sequencing (500 bp approx.) of the latter with no differences being detected. Primer extension (results not shown) indicated that the longer clone was full length. The cDNA encodes one major open reading frame which extends from a start codon at position 136 to a stop codon at position 1,102 encoding a protein of 322 amino acids. There are 135 bp of untranslated 5′ sequences and 230 bases of 3′ sequence terminated by a polyA sequence of 7 nucleotides (not shown) with a potential polyadenylation signal, GATAAA, 29 bp upstream (1,303–1,308).

*Corresponding author. Fax: (44) (382) 562426.

SSDI 0014-5793(94)01105-2

```
aaaacccttcactcgcttatctgcgcgatccctagttaacgccatggccggaacaccgagaacttctccaccgacgctac   80

tgccgacgccgcaacgccaacggaggatgtcaagtttgggtttcagcgttcggagatgtatcagagtaaacttgcaggca  160
                                                          M  Y  Q  S  K  L  A  G  T    9

ctgcgacctcctatgatcgtcatctgttcctctgttacaagtctcacgagacttgtcctgctcgccttgaagcttctgac  240
  A  T  S  Y  D  R  H  L  F  L  C  Y  K  S  H  E  T  C  P  A  R  L  E  A  S  D     35

tccgatctgcttcccaagtcattctctgctgctcttaaagctcgcaaagatgatatcaaaatcaagactcttttgacaat  320
S  D  L  L  P  K  S  F  S  A  A  L  K  A  R  K  D  D  I  K  I  K  T  L  L  T  I    62

atgcgaagtacgtgatgacatggaagtatcagaaggagatattctgatttttcctgaaatgatcaaatacagggatttga  400
 C  E  V  R  D  D  M  E  V  S  E  G  D  I  L  I  F  P  E  M  I  K  Y  R  D  L  K   89

aggagtcagatgtggatgcatttgttgatgatgtgcttgtcaatggcaatccttggagctctggattgcaggagtcactg  480
  E  S  D  V  D  A  F  V  D  D  V  L  V  N  G  N  P  W  S  S  G  L  Q  E  S  L    115

agtggttcttatgtgtttgtctgtgcccacaatcttcgtgatcgaagatgtggtgtttgcggacctattctaattgagga  560
S  G  S  Y  V  F  V  C  A  H  N  L  R  D  R  R  C  G  V  C  G  P  I  L  I  E  E   142

atttagcaagctcattgagtccaagggcttgaaagacaaagttcgtgtggcagcttgttctcatattggtggccacaagt  640
 F  S  K  L  I  E  S  K  G  L  K  D  K  V  R  V  A  A  C  S  H  I  G  G  H  K  Y  169

atgctggtaatgttataatcttcagctcaggaaaggacggggatattgttggccactggtatggctatgttacgccaagt  720
  A  G  N  V  I  I  F  S  S  G  K  D  G  D  I  V  G  H  W  Y  G  Y  V  T  P  S    195

gatgtacctgctttgcttgatgagcatattggagagggaaaagtcattgaacgactttggaggggccaaatgggacaata  800
D  V  P  A  L  L  D  E  H  I  G  E  G  K  V  I  E  R  L  W  R  G  Q  M  G  Q  Y   222

tgagaaggtaactgataaagtggatgaacagaaggttcctgaagtaaccaatgaagaaaagaagcctcttgaaaatggaa  880
 E  K  V  T  D  K  V  D  E  Q  K  V  P  E  V  T  N  E  E  K  K  P  L  E  N  G  S  249

gtcaggagagtagtgtaactagtttcagctgttgccaaggtgctgcgggagtttcttgttgtagagatgcaagtgctgag  960
  Q  E  S  S  V  T  S  F  S  C  C  Q  G  A  A  G  V  S  C  C  R  D  A  S  A  E    275

caggaagagaacaagaaagggcagggaacagtttcaaactggttcggcaaatgggagcagcgtgaaatactcgcacgtgt 1040
Q  E  E  N  K  K  G  Q  G  T  V  S  N  W  F  G  K  W  E  Q  R  E  I  L  A  R  V   302

aggtgtggttggagcagtggcagttgtcgctgtggcttatggcttttacaagaagtcacggtgaagcatgttaactcagg 1120
 G  V  V  G  A  V  A  V  V  A  V  A  Y  G  F  Y  K  K  S  R  *                    322

acaaactttgtgctcaagagttttcgggtatctaggttttgtagaaaggtaataatgttcgagataatctcggacactga 1200

aaatgcatacatatttggtatattgagagaattaaatgcgatttgtaattcgactatggagatgagctcctttgctcgtt 1280

cccagtaagaactagaggggctgataaatatgaagtgccaactgattgctac    1332
```

Fig. 1. Nucleotide and derived amino acid sequence from pEF141. Underlining indicates position of sequences used for primer design for RT-PCR (see Fig. 2b)

3.2. Genomic organisation and expression

The organisation of the gene corresponding to this cDNA was examined by Southern hybridisation (Fig. 2a). The discrete number of hybridising bands observed from this tetraploid potato variety is indicative of a low copy number for this gene. Segregation analysis on a diploid potato mapping population [10] allowed a single locus to be mapped to potato chromosome II between CP116 and CP70(a) (C. Gebhardt, personal communication). Expression of the gene was detected at similar levels in all potato tissues examined by RT-PCR (Fig. 2b), yielding the expected product of 229 bp with no product in control reactions omitting reverse transcriptase (confirming RNA as the source of the expression product).

3.3. Sequence analysis

Both the nucleotide and derived amino acid sequences of pEF141 have been used to search databases for similar sequences using a variety of programs. A highly significant similarity (see Fig. 3a) was detected between the central amino acid domain (residues 120–214) and the 103 amino acid ferredoxin from *Clostridium pasteurianum* [11] (28% identity over 100 amino acids with an optimised FASTA [12] score of 118, with the COOH-proximal region (amino acids 186–214) scoring 80 ($P = 0.0024$) in BLAST [13]). Neither this central domain, nor the clostridial ferredoxin, show significant similarity to other proteins or ferredoxins which include the 'plant' type of the chloroplast membrane and the 'Rieske'-type found in mitochondria [14]. There are three classes of ferredoxin which functions as an electron carrier. These are distinct in their physical, chemical and magnetic properties but all contain [Fe-S] groups. The clostridial ferredoxin contains a [2Fe-2S] cluster. Bacterial ring-hydroxylating dioxygenases also contain ferredoxin within the enzyme complex, either as a separate protein, e.g benzene, toluene, pyrazon and napthalene dioxygenases, or as a domain within a larger protein which also includes a flavin, e.g. benzoate 1,2-, phthalate and 4-chlorophenylacetate-3,4-dioxygenases (for review see [15]). Both 'Rieske' and 'plant' type ferredoxin [Fe-S] domains have been co-opted into these larger proteins with a hydrolase function. In other proteins [Fe-S] domains may have the additional or alternative function of substrate binding. They are found in diverse eukaryotic proteins, including aconitase [16] and dihydroxyacid dehydratase [17] in plants. The latter enzyme in bacteria [18] also contains [Fe-S] moieties, as do glutamine PRP amidotransferase [19], fumarate reductase [20] and endonuclease III [21]. However no protein including a domain similar to the clostridial ferredoxin has previously been described.

Sequence searches were next conducted with the nucleotide sequence 5′ to that encoding the ferredoxin-like domain, and its derived amino acid sequence. A second highly significant similarity was detected at the nucleotide level. Nucleotides 86–314 show 61% identity over 228 bp with nucleotides 4,410–4,642 of a genomic clone from *Arabidopsis thaliana* [22] (cm_pl:s50281 in the EMBL database). This sequence lies

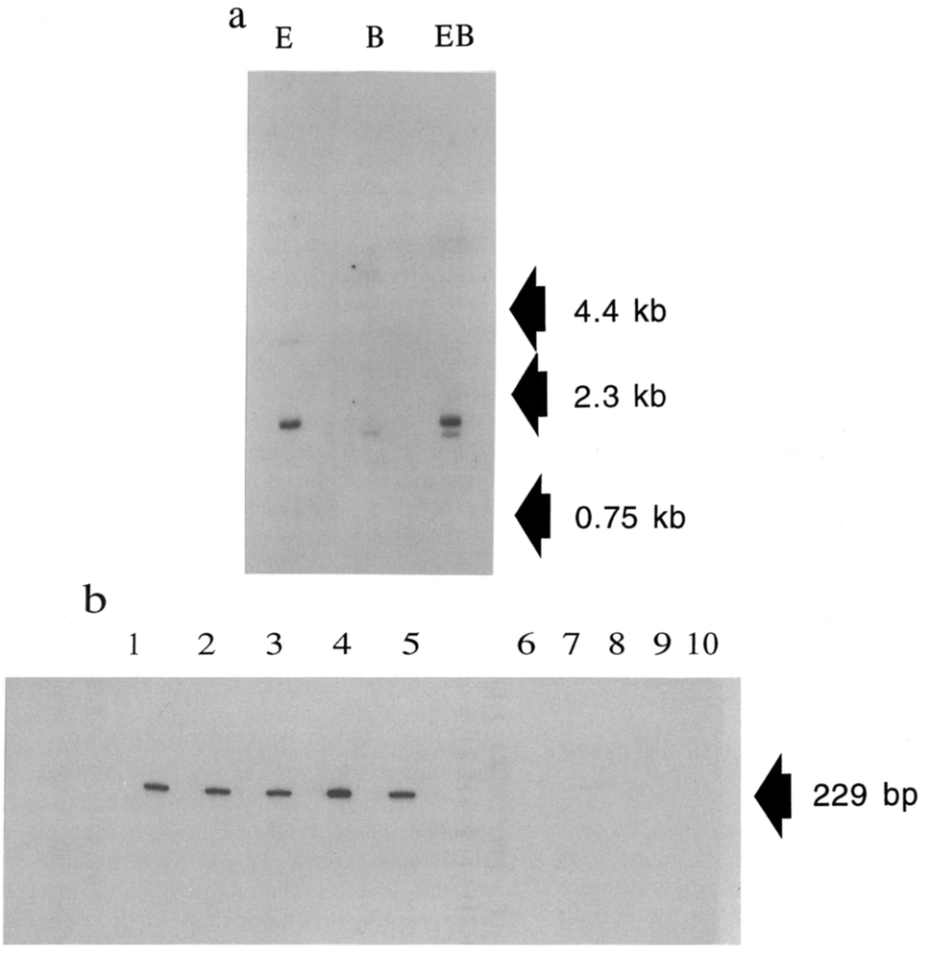

Fig. 2. (a) Southern blot of potato (*Solanum tuberosum* cv. Saturna) genomic DNA digested with *Eco*RI (E), *Bam*HI (B), and both enzymes in combination (EB) and probed with the 1,332 bp insert of plasmid pEF141. Size in kb is estimated from bands representing fragments on the same gel hybridising to a control probe. (b) RT-PCR analysis of transcripts. Oligonucleotides for RT-PCR were nt 247–272 and nt 476–451 (numbered as in Fig. 1), RNA samples were (1) source leaf, (2) sink leaf, (3) stem, (4) root, and (5) tuber. In samples 6–10, reverse transcriptase was omitted from the reaction. Products were sized by comparison (after longer exposure) to the DNA sequencing markers of known size run between lanes 5 and 6.

downstream of the Atpk7 gene, encoding a protein kinase homologue, described from this clone. No open reading frame was ascribed to this region of the clone but nucleotides 4,459–4,621 encode a derived amino acid sequence of 54 residues with 46% identity to the amino-terminal of the derived amino acid sequence from pEF141 (see Fig. 3b). This similarity is terminated by stop codons in all three frames of the derived amino acid sequence from the *Arabidopsis* gene in a region very close to the end (nucleotide 4,719) of this clone. Partial sequencing of a genomic clone corresponding to the potato cDNA (unpublished results) indicates the presence of an intron which, if also present in the *Arabidopsis* gene, would extend from position 4,622 to the end of the *Arabidopsis* clone. The potato genomic clone also carries a potato homologue of the Atpk7 gene. These data are consistent with this region of the *Arabidopsis* genome encoding the equivalent of the potato cDNA, but neither sequence is highly related to any other sequence in the databanks. The nucleotide sequence 3′ to the ferredoxin-like domain

```
a.

ef 120  VFVC.AHNLRDRRCGVC....GPILIEEFSKLIESKGLKDKVRVAACSHIGGHKYAGNVIIFSSGKDGDIVGHWYGYVTPSDVPALLDEHIGEGKVIERL 214
fe   8  IFVCTSCRLNGKQQGFCYSKNSVEIVETFMEELDSRDLSSEVMVNNTGCFGICSQGPIVVVYPEGV.......WYGNVTADDVEEIVESHIENGEVVKRL 100

b.

ef   1  MYQSKLAGTATSYDRHLFLCYKSHETCPARLEASDSDLLPKSFSAALKARKDDIKIK 57
At   1  MYSGTLAGSVGPYGRHVFLCYKSHET...WLPRVETEGLPQRFAKLFKDRKADFAVE 54
```

Fig. 3. Comparison of domains from the derived amino acid sequence of pEF141 (ef) to database sequences (a) the derived amino acid sequence from the *Clostridium pasteurianum* ferredoxin gene [11]; (b) derived amino acid sequence from an *Arabidopsis thaliana* genomic clone ([22]; em_pl:s50281 in the EMBL database).

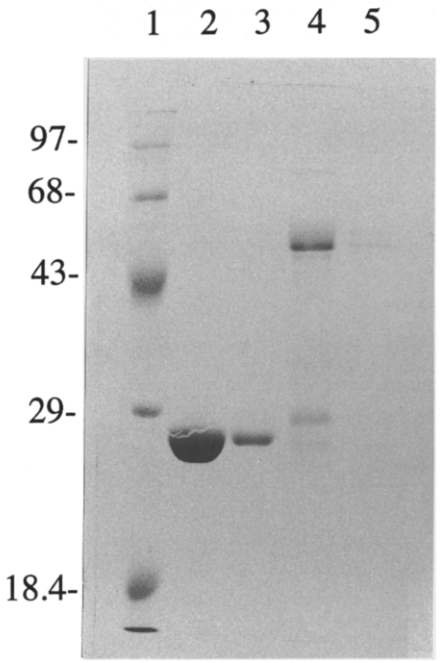

Fig. 4. SDS-polyacrylamide gel electrophoresis of affinity prepared expression proteins. Lane 1: molecular weight markers (size in kDa); lanes 2 and 3: glutathione-*S*-transferase (GST) expressed from pGEX2T; lanes 4 and 5: fusion protein from partial ORF of pEF141 cloned into pGEX2T.

encodes a polypeptide sequence enriched for the amino acids lysine and glutamic acid. Database searches using this sequence select only proteins of diverse function which are rich in these two amino acids.

3.4. Expression clone

To examine function of the protein encoded by the cDNA cloned in pEF141, the majority of the ORF (from the *Hin*dIII at nt 231 to the end of the cDNA) was cloned in frame into the expression vector pGEX2T. The fusion protein induced by IPTG was prepared by affinity chromatography (see Fig. 4). The fusion protein (lanes 4 and 5) has an estimated molecular weight of 57 kDa and reacts specifically with the antibody used for the selection of clones from the cDNA library. Using 0.3 mg of purified protein per assay, neither iron nor inorganic sulfide could be detected in the fusion protein, which also lacked any spectroscopic properties similar to those characteristic of the [2Fe-S] chromophore of clostridial ferredoxin [23]. Supplementing the bacterial growth medium with 100 μM $FeSO_4$ had no effect on the properties of the fusion protein. Either the sequence similarity to clostridial ferredoxin noted above is not consequent from functional similarity between these proteins or assembly of the Fe-S chromophore, which may function as a dimer in solution, is not achieved in the fusion protein in *E. coli*. While *E. coli* has been successfully used for the heterologous expression of several ferredoxins, generally low and highly variable yields of holoferredoxin have been reported, resulting from problems of stability or efficiency of insertion of the Fe-S cluster (see [23] for a discussion). Purification of the native protein from potato will be required to determine its status as an Fe-S protein.

We next assayed for enzyme activity against a variety of substrates, chosen on the basis of the sucrolytic activity of the protein from which the original antibody was raised. Using sucrose as substrate a sucrase activity equivalent to 1.84 ng glucose/h/μg protein was detected for the fusion protein, with a K_m for sucrose of 2.01×10^2 mM. This is a low level of activity and substrate affinity by comparison to invertases previously characterised in potato which exhibit a range of K_m for sucrose from 2.4 to 16 mM [6,24], but may reflect aberrant enzyme assembly resulting from factors such as constraints placed by the addition of the glutathione-*S*-transferase sequences, inefficient incorporation of [Fe-S] clusters into the fusion protein or lack of glycosylation. Alternatively it may indicate that sucrose is not the primary physiological substrate for this enzyme. Activity was dependent on the presence of dithiothreitol, again raising the possibility of redox control of the enzyme, achieved by the supplied reducing agent in the absence of correct assembly of a native holoenzyme including [Fe-S] clusters. No activity was detected using a range of other sugars (raffinose, stachyose, trehalose, lactose and maltose) as substrate and control purifications of the glutathione-*S*-transferase polypeptide expressed from pGEX2T (Fig. 4) showed no activity against sucrose. The host *E. coli* (DH5α) used for fusion protein preparation lacks any sucrase activity. While the ability of the enzyme produced in *E. coli* to cleave sucrose is consistent with the fact that its cDNA was cloned using an antibody raised against a protein preparation with sucrolytic activity, a more detailed study of its catalytic properties will require purification of the native enzyme from potato. It will be interesting to determine whether the enzyme described here is involved in sucrose cleavage under physiological conditions – it is clearly different from the two enzymes, invertase and sucrose synthase, to which this function has previously been ascribed. A full description of its wider role in plant cell metabolism will require manipulation of its expression in transgenic plants.

Acknowledgements: We thank N. Harris and E. Cuthbert for technical support in the assay of invertase activity. This research was supported by ECSA Research Ltd. and the EC under the ECLAIR programme. HVD and RW are funded by the Scottish Office Agriculture and Fisheries Department. This work benefited from use of the SEQNET facility at SERC, Daresbury Laboratory, UK.

References

[1] Smith, C.J. (1993) In: Plant biochemistry and molecular biology. (Lea P.J. and Leegood R.C. eds.), pp. 73–111, Wiley and Sons Ltd., Chichester.
[2] Copeland, L. (1990) Methods Plant Biochem. 3, 73–85.
[3] Pollock, C.J. (1986) New Phytologist 104, 1–24.
[4] Hedley, P.E., Machray, G.C., Davies, H.V., Burch, L.R. and Waugh, R. (1993) Plant Mol. Biol. 22, 917–922.
[5] Hedley, P.E., Machray, G.C., Davies H.V., Burch L.R. and Waugh R. (1994) Gene, in press.
[6] Burch, L.R., Davies, H.V., Cuthbert, E.M., Machray, G.C., Hedley, P. and Waugh R. (1991) Phytochemistry 31, 1901–1904.
[7] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[8] Simpson, C.G., Sinibaldi, R. and Brown, J.W.S. (1992) The Plant J. 2, 835–836.
[9] Lovenberg, W., Buchanan, B.B. and Rabinowitz, J.C. (1963) J. Biol. Chem. 238, 3899–3912.
[10] Barone, A., Ritter, E., Schachtschabel, U., Debener, T., Salamini, F., Gebhardt, C. (1990) Mol. Gen. Genet. 224, 177–182.
[11] Meyer J. (1993) Biochim. Biophys. Acta 1174, 108–110.
[12] Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.

[13] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403–410.
[14] Matsubara, H. and Saeki, K. (1992) Advances in Inorganic Chemistry 38, 223–280.
[15] Mason J.R. and Cammack R. (1992) Annu. Rev. Microbiol. 46, 277–305.
[16] Brouquisse, R., Nishimura, M., Gaillard, J. and Douce, R. (1987) Plant Physiol. 84, 1402–1407.
[17] Flint, D.H. and Emptage, M.H. (1988) J. Biol. Chem. 263, 3558–3564.
[18] Flint, D.H., Emptage, M.H., Finnegan, M.G., Fu W.G. and Johnson, M.K. (1993) J. Biol. Chem. 268, 14732–14742.
[19] Vollmer, S.J., Switzer, R.L. and Debrunner, P.G. (1983) J. Biol. Chem. 258, 14284–14293.
[20] Johnson, M.K., Kowal, E.T., Morningstar, J.E., Oliver, M.E., Whittaker, K., Gunsalas, R.P., Ackrell, B.A.C. and Cecchini, G. (1988) J. Biol. Chem. 263, 14738–14742.
[21] Cunningham, R.P., Asahara, H., Bank, J.F., Scholes, C.P., Salerno, J.C., Surerus, K., Munck, E., McCracken, J., Peisach, J. and Emptage, M.H. (1989) Biochemistry 28, 4450–4455.
[22] Hayashida, N., Mizoguchi, T., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1992) Gene 121, 325–330.
[23] Fujinaga, J. and Meyer, J. (1993) Biochem. Biophys. Res. Commun. 192, 1115–1122.
[24] Bracho, G.E. and Whitaker, J.R. (1990) Plant Physiol. 92, 386–394.