Volume 263, number 2, 274–278 FEBS 08348 April 1990

Polypeptide composition of higher plant photosystem I complex

Identification of *psa*I, *psa*J and *psa*K gene products

Masahiko Ikeuchi, Atsushi Hirano°, Tetsuo Hiyama+ and Yorinao Inoue

Solar Energy Research Group, The Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-01, °Engineering Research Center, Tokyo Electric Power Company, Chofu, Tokyo 182, and *+Department of Biochemistry, Faculty of Science, Saitama University, Urawa, Saitama 338, Japan*

Received 10 February 1990

High resolution gel electrophoresis of the native photosystem I complex retaining light-harvesting chlorophyll complex revealed the presence of three low-molecular-mass proteins of 7, 4.1 and 3.9 kDa in spinach, and 6.8, 4.4 and 4.1 kDa in pea, in addition to the other well-characterized higher-molecular-mass components. Upon further detergent treatment to deplete light-harvesting chlorophyll complex, the 7 kDa and 4.1 kDa proteins were removed from the photosystem I core complex of spinach, while the 3.9 kDa protein was retained. N-terminal sequencing demonstrated that the 4.1 kDa proteins from both spinach and pea correspond to the gene product of ORF42/44 in chloroplast genome of liverwort and higher plants, which was previously hypothesized as a photosystem I gene (*psa*J) based on sequence homology with the cyanobacterial photosystem I component of 4.1 kDa [(1989) FEBS Lett. 253, 257–263]. N-terminal sequence of the spinach 3.9 kDa and pea 4.4 kDa proteins fitted with chloroplast ORF36/40 (*psa*I) although no homologue has been found in cyanobacteria. The spinach 7 kDa and pea 6.8 kDa proteins correspond to the nuclear-encoded *psa*K product and significantly matched with the N-terminal sequence of the cyanobacterial 6.5 kDa subunit. The evolutional conservation of the *psa*J and *psa*K seems to suggest their intrinsic role(s) in photosystem I.

Photosystem I; Gene, *psa*I; Gene, *psa*J; Gene, *psa*K; (Spinach; Pea)

1. INTRODUCTION

PSI drives electron flow from cytochrome b_6/f to ferredoxin/NADP in thylakoid membranes of both cyanobacteria and green plants. PSI can be extracted by various detergent treatments as 'core' complex, which consists of P700-carrying reaction center and several colorless core subunits [1]. 'Native' PSI complex which additionally retains LHCI can be isolated from green plants by a milder detergent treatment, while cyanobacterial PSI has no LHCI-like antenna. Until recently the subunit proteins found in these various types of PSI complexes from both green plants and cyanobacteria have been classified in three categories as follows: (i) two chlorophyll *a*-binding subunits of about 80 kDa, which carry P700, acceptors A_0 and A_1, and Fe-S center X, (ii) 5–6 core subunits ranging from 8 to 22 kDa which include Fe-S center A/B protein, plastocyanin-docking protein and ferredoxin-docking protein, (iii) 4–5 LHCI subunits of green plants ranging from 15 to 25 kDa which bind antenna chlorophyll *a* and *b*. Sequencing of these proteins and/or their corresponding genes have registered two genes, *psa*A and *psa*B, for the P700-carrying proteins [2,3], 6 genes, *psa*C to *psa*H, for the core subunits [4–17] and at least three types of *cab* genes for the LHCI subunits [18–20]. In green plants, *psa*A, *psa*B and *psa*C are encoded by chloroplast genome while the others are encoded by nuclear genome. Sequence comparison indicates that the two P700-carrying proteins and most of the core subunits are more or less conserved between cyanobacteria and green plants [3,13,17], although LHCI is absent in cyanobacteria.

In addition to these components, we recently found three new low-molecular-mass components of 4.1 kDa, 5 kDa and 6.5 kDa in cyanobacterial PSI complex and reported their partial amino acid sequences [17]. Interestingly, the sequence of this cyanobacterial 4.1 kDa component corresponded to ORF42/44 of higher plant chloroplast DNA, although its product had not yet been found in plants. Scheller et al. [21] reported the presence of two small proteins below 4 kDa in higher plant PSI complex. These prompted us to search for the product(s) of ORF42/44 in higher plant PSI complex as well. Here we report the presence of three low-molecular-mass proteins in spinach and pea PSI complexes and their unambiguous identification by N-terminal sequencing.

Correspondence address: M. Ikeuchi, Solar Energy Research Group, The Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-01, Japan

Abbreviations: CPI, chlorophyll-protein complex I of photosystem I reaction center; LHCI, light-harvesting chlorophyll complex associated with photosystem I; PSI, photosystem I; SDS-PAGE, sodium dodecylsulfate-polyacrylamide gel electrophoresis

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

2. MATERIALS AND METHODS

Native PSI complex retaining LHCI ('PSI-200') was isolated from spinach (*Spinacia oleracea*) and pea (*Pisum sativum*) thylakoids according to Mullet et al. [22]. Thylakoid membranes were washed once with 5 mM EDTA and 50 mM sucrose, pH 7.8, resuspended in distilled water and then treated with 1% (w/v) Triton X-100 at 0.8 mg Chl/ml. Solubilized materials were fractionated by centrifugation at 131000 × *g* for 15 h on a density gradient of 0.1–1.0 M sucrose containing 0.02% (w/v) Triton X-100 with a 2 M sucrose cushion containing 0.02% (w/v) Triton X-100. The native PSI complex with chlorophyll *a* to *b* ratio of about 6 was obtained at the interface between 1 and 2 M sucrose layers. To deplete LHCI, the native complex was further treated with 1% (w/v) Triton X-100 at 0.2 mg Chl/ml and fractionated by centrifugation at 131000 × *g* for 9 h on a density gradient of 0.1–1.0 M sucrose containing 0.35% (w/v) Triton X-100. The resulting PSI core complex with chlorophyll *a* to *b* ratio higher than 10 was recovered from the middle part of the gradient.

High resolution SDS-PAGE with 7.5 M urea and a 16–22% (w/v) acrylamide gradient was done according to [23]. For reproducible resolution of low-molecular-mass proteins, the pH of resolving gel buffer (Tris-HCl) was adjusted to pH 8.4 instead of pH 8.8. Proteins in a gel were transferred onto a polyvinylidene difluoride membrane (Immobilon, 0.45 μm, Millipore) following [24,25]. Transferred proteins were stained with Amido Black 10B (Bio-Rad), cut out and subjected to amino acid sequencing with a protein sequencer (model 477A, Applied Biosystems). When indicated, transferred proteins were treated with 0.6 N HCl for 24 h to release N-terminal block before sequencing.

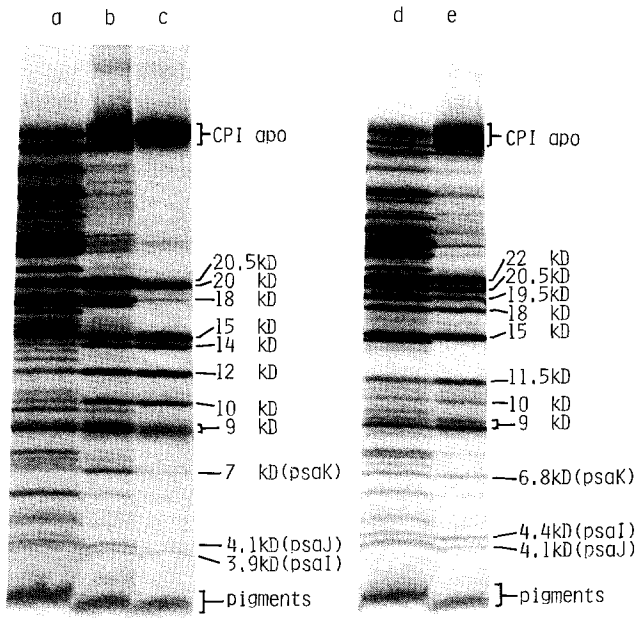

Fig. 1. Polypeptide composition of PSI complexes from spinach and pea. Lanes a–c, spinach; d, e, pea. Lanes a, d, thylakoid membranes; b, e, native PSI complex; c, PSI core complex.

3. RESULTS AND DISCUSSION

High resolution SDS-PAGE revealed the presence of three components below the 9 kDa doublet band in both spinach and pea PSI complexes in addition to at least 9 bands ranging from 9 kDa to 22 kDa (Fig. 1). Molecular masses of the three low-molecular-mass proteins are estimated as 7 kDa, 4.1 kDa and 3.9 kDa for spinach and 6.8 kDa, 4.4 kDa and 4.1 kDa for pea, based on the relationship experimentally determined between mobilities and molecular masses of PSII intrinsic proteins [23]. All these bands originating from the PSI complex are found as discrete bands in thylakoid membranes as well, suggesting that these are not artifacts due to proteolysis during isolation but are intrinsic components of PSI. Further detergent treatment not only depleted LHCI but also extracted the 7 kDa and 4.1 kDa proteins from the spinach core complex, while leaving the 3.9 kDa protein still associated with the core complex (Fig. 1, lane c). Other PSI subunits of spinach in higher-molecular-mass region were identified by protein sequencing as follows: (i) the 20 kDa, 15 kDa, 12 kDa and 10 kDa proteins are gene products of *psa*D, *psa*F, *psa*E and *psa*H, respectively (data now shown), (ii) the 9 kDa doublet bands contain gene products of *psa*C and *psa*G, (iii) several new sequences were additionally obtained from the 20 kDa, 18 kDa and 9 kDa bands, and the 20 kDa and 18 kDa sequences seem to correspond to type I and II LHCI, respectively [26], (iv) the 20.5 kDa and 14 kDa proteins have blocked N-termini. Further detergent treatment of the native PSI complex removed the *psa*G product, LHCI subunits and most of unknown subunits from the core complex except for the 14 kDa protein. Separation profile of the pea PSI complex is similar to that of spinach, although N-terminal sequencing of the pea subunits remains not enough for their unambiguous assignment.

N-terminal sequencing of both spinach and pea 4.1 kDa proteins provided a pair of signals at each sequencing cycle, one strong signal and one weak signal. The strong signal was obtained only after HCl treatment, indicative of N-terminal block of this protein. Both spinach and pea sequences compiled from the strong signals (Fig. 2) clearly corresponded to the N-terminal sequences deduced from ORF42/44 of plant chloroplast DNA [27–29]. The *Synechococcus* 4.1 kDa protein, whose N-terminus was also blocked [17], was largely homologous but somewhat more divergent in its N-terminal region. Based on the correspondence of N-terminal sequence of the *Synechococcus* 4.1 kDa protein to the DNA sequence of plant ORF42/44, we tentatively designated this ORF *psa*J [30]. The *psa*J gene is located downstream of a ribosomal subunit gene (*rpl*33) of chloroplast DNA in tobacco and rice [27,28] as well as in liverwort [29]. Our direct sequencing in this study unambiguously confirmed the association of this protein with the native PSI complex of spinach, although it was not retained in the PSI core complex (Fig. 1). Based on the amino acid sequences deduced from tobacco ORF44 and liverwort ORF42, a possible membrane-spanning domain was found (Fig. 2, also see [17]). Hydrophobicity of the domain and a

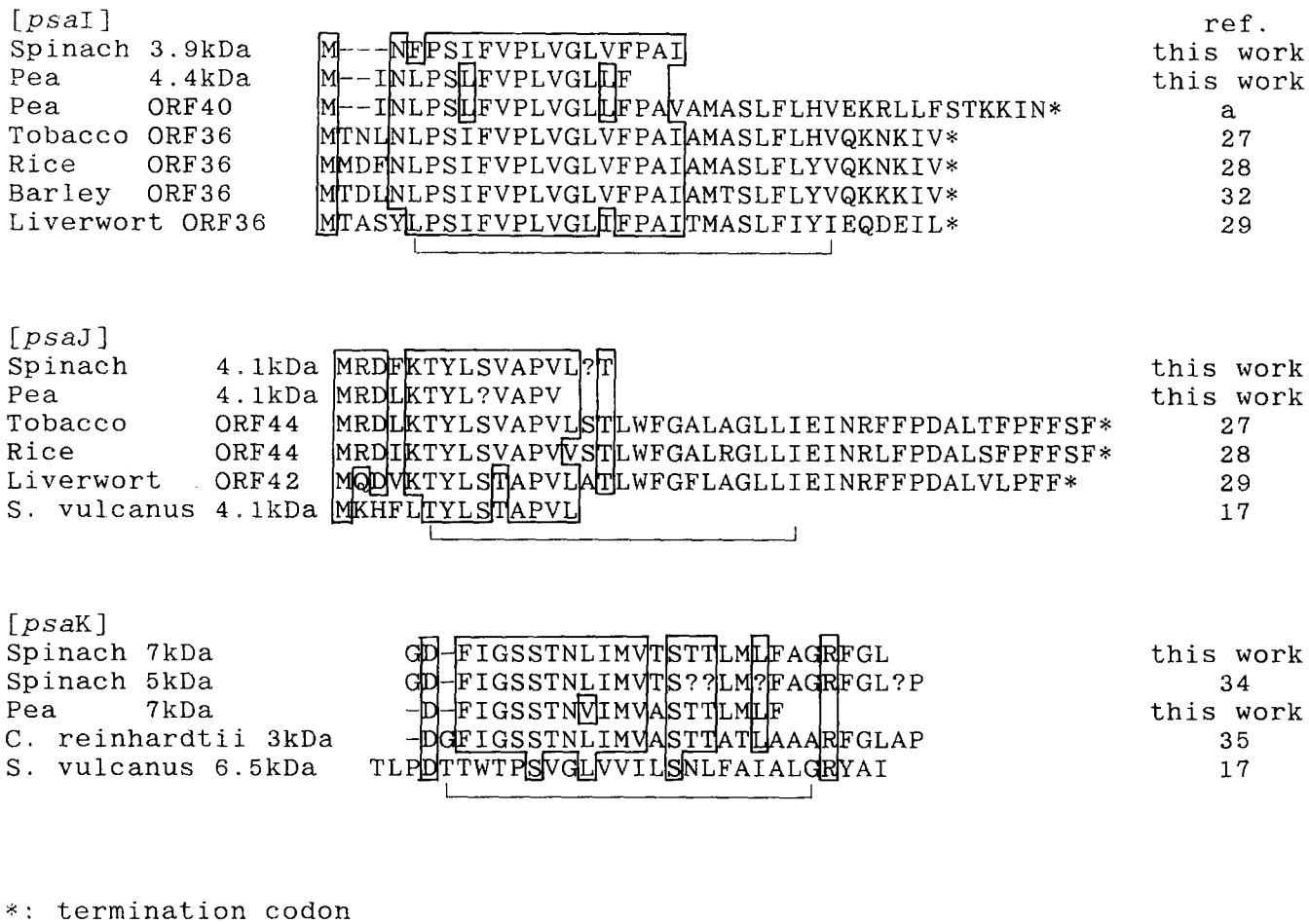

Fig. 2. N-terminal sequences of spinach 3.9 kDa, 4.1 kDa and 7 kDa proteins and pea 4.4 kDa, 4.1 kDa and 6.8 kDa proteins aligned with the known sequences of ORF36/40 (*psa*I), ORF40/42 (*psa*J) and *psa*K. Conserved amino acid residues are boxed. When sequences are deduced from genes, full-length sequences are shown, with the exception of the too large *Chlamydomonas* 8.4 kDa protein. *, Termination codon; a, Nagano and Sasaki, personal communication.

preceding charged Lys residue were both conserved in spinach and pea, suggestive of its membrane-spanning properties. This agrees with the fact that the 4.1 kDa protein can be extracted by chloroform/methanol (results not shown). The satellite sequence compiled from the weak signals of the spinach 4.1 kDa band corresponded to a nuclear-encoded component of PSII core complex [24], while the similar satellite sequence from the pea band corresponded to the *psb*K protein of PSII. Probably these are due to cross-contamination of PSII complex in the PSI fraction.

N-terminal sequence of the pea 4.4 kDa protein, which was obtained only after HCl treatment, completely matched with the deduced sequence of pea ORF40, downstream of *zfp*A in chloroplast DNA (Fig. 2, see also [31]). The apparent molecular mass of the pea protein is consistent with the calculated molecular mass (4469 Da) of the deduced product. Homologous sequence was obtained from the spinach 3.9 kDa protein after HCl treatment. In a similar position of chloroplast DNA of tobacco, rice and liverwort, ORF36 is found to be homologous to these N-terminal sequences (Fig. 2). Calculated molecular masses of the ORF36 of these plants are 3900–4000 Da, which are close to the apparent molecular mass of the spinach protein (4.1 kDa). When these sequences and ORF36/40 of higher plants and liverwort are compared, high conservation is found only in internal relatively hydrophobic segment of 24 residues, marked in Fig. 2, involving only a few conservative replacements. This segment may span the thylakoid membrane since this protein was extractable by chloroform/methanol (not shown). In contrast, N-terminal and C-terminal regions of this component are highly varied: two and three residues are missing in the N-terminal region of the pea and spinach proteins, respectively, as shown by the alignment in Fig. 2, and in addition, the pea sequence has a C-terminal extension of 6 amino acid residues. This C-terminal extension seems to account for the difference in apparent

molecular mass of this protein between pea and spinach.

Recently Scheller et al. [32] reported the homologous N-terminal sequence for the 4 kDa component of barley PSI core complex and designated the corresponding ORF36 of chloroplast DNA as *psa*I. Based on the weak sequence homology they claimed between this protein and transmembrane helix E of PSII D2 protein, they proposed that this small protein together with other hypothetical small components might form the reaction center of PSI instead of CPIa and CPIb subunits. However, this is unlikely because we found that the spinach 4.1 kDa protein or any other low-molecular-mass component are not retained in active PSI reaction center complex consisting of only CPIa and CPIb (Hiyama, T., Ikeuchi, M. and Inoue, Y., unpublished data) which was prepared by brief exposure of the core complex to SDS [33]. Furthermore, we cannot assign any components in *Synechococcus* PSI core complex as a homologue of the spinach 3.9 kDa protein [17]. Thus, we may exclude the possibility that the spinach 3.9 kDa protein plays a central role in the PSI reaction center.

N-terminal sequences of the spinach 7 kDa and pea 6.8 kDa proteins, which were obtained without HCl treatment, are highly homologous to each other (Fig. 2). Major difference between the two is found in the first residue; N-terminal Gly residue of the spinach protein is missing in the pea protein. The spinach sequence is identical to that of spinach '5 kDa' protein, which was recently detected in heat-treated PSI complex depleted of Fe-S center A/B but retaining Fe-S center X [34]. These proteins are also highly homologous to an 8.4 kDa protein of *Chlamydomonas* PSI complex, whose gene sequence was recently determined and designated *psa*K [35]. In all these sequences there is one conserved hydrophobic domain as is marked in Fig. 2. The charged residues surrounding this hydrophobic domain are also conserved among them. These suggest that the hydrophobic domain may span the membrane and the charged residues may function as 'stop transfer'. In agreement with this, *Chlamydomonas* 8.4 kDa protein is extracted by chloroform/methanol (denoted P37 in [36]). It may also be of note that these sequences are weakly homologous to that of the *Synechococcus* 6.5 kDa protein (Fig. 2). Interestingly, the membrane-spanning properties of these plant proteins, as suggested by their hydrophobic domain and the flanking charged residues, are preserved in the *Synechococcus* protein as well. These similarities may imply that the *psa*K product plays a common role in PSI of both cyanobacteria and green plants. In addition, Hoshina et al. [34] demonstrated by heat treatment of their PSI complex that this protein is directly associated with the PSI reaction center. However, the results in this study show that this protein is not recovered into PSI core complex when prepared by detergent treatment. This clearly excludes the possibility of the involvement of this component in PSI activity. These results together with its hydrophobic nature may rather suggest a structural role of the *psa*K product in PSI, i.e., stabilization of the reaction center.

Acknowledgements: We thank Drs Y. Sasaki and Y. Nagano (Kyoto University, Japan) for kindly allowing us to use their unpublished gene sequence of pea ORF40 (*psa*I). We also thank Dr D. Bryant (Pennsylvania State University, USA) for his valuable suggestion about the sequence homology between the *Synechococcus* 6.5 kDa protein and the *Chlamydomonas* 8.4 kDa protein. We thank Dr S. Hoshina (Kanazawa University, Japan), Dr L.-G. Franzén (Université de Genève, Switzerland) and Dr B.L. Møller (Royal Veterinary and Agricultural University, Denmark) for providing us their preprints. We thank Dr Y. Narahashi (RIKEN) for her helpful instructions in protein sequencing and Dr H. Koike (RIKEN) for his encouraging discussion. This work was supported by a Grant on Solar Energy Conversion by Means of Photosynthesis awarded to the Institute of Physical and Chemical Research (RIKEN) by the Science and Technology Agency (STA) of Japan, and partly by a Grant on Frontier Research Program at RIKEN awarded by STA and an MESC Grant-in-Aid for Cooperative Research, No. 01302064.

REFERENCES

[1] Golbeck, J.H. (1988) Biochim. Biophys. Acta 895, 167–204.
[2] Fish, E., Kuck, U. and Bogorad, L. (1985) J. Biol. Chem. 261, 8134–8139.
[3] Cantel, A. and Bryant, D.A. (1987) Plant Mol. Biol. 9, 453–468.
[4] Hayashida, N., Matsubayashi, T., Shinozuka, K., Sugiura, M., Inoue, K. and Hiyama, T. (1987) Curr. Genet. 12, 247–256.
[5] Oh-oka, H., Takahashi, Y., Wada, K., Matsubara, H., Ohyama, K. and Ozeki, H. (1987) FEBS Lett. 218, 52–54.
[6] Høj, P.B., Svendsen, I., Scheller, H.V. and Møller, B.L. (1987) J. Biol. Chem. 262, 12676–12684.
[7] Lagoutte, B. (1988) FEBS Lett. 232, 275–280.
[8] Hoffman, N.E., Pichersky, E., Malik, K.K. and Cashmore, A.R. (1988) Plant Mol. Biol. 10, 435–445.
[9] Steppuhn, J., Hermans, J., Nechushtai, R., Ljungberg, U., Thümmler, F., Lottspeich, F. and Herrmann, R.G. (1988) FEBS Lett. 237, 218–224.
[10] Okkels, J.S., Jeppsen, L.B., Hønberg, L.S., Lehmbeck, J., Scheller, H.V., Brandt, P., Høyer-Hansen, G., Stummann, B., Henningsen, K.W., Von Wettstein, D. and Møller, B.L. (1988) FEBS Lett. 237, 108–112.
[11] Münch, S., Ljunberg, U., Steppuhn, J., Schneiderbauer, A., Nechushtai, R., Beyreuther, K. and Herrmann, R.G. (1988) Curr. Genet. 14, 511–518.
[12] Alhadeff, M., Lundell, D.J. and Glazer, A.N. (1988) Arch. Microbiol. 150, 482–488.
[13] Reilly, P., Hulmes, J.D., Pan, Y.-C.E. and Nelson, N. (1988) J. Biol. Chem. 263, 17658–17662.
[14] Steppuhn, J., Hermans, J., Nechushtai, R., Ljungberg, U., Thümmler, F., Lottspeich, F. and Herrmann, R.G. (1989) Curr. Genet. 16, 99–108.
[15] Okkels, J.S., Scheller, H.V., Jepsen, L.B. and Møller, B.L. (1989) FEBS Lett. 250, 525–579.
[16] Franzén, L.-G., Frank, G., Zuber, H. and Rochaix, J.-D. (1989) Plant Mol. Biol. 12, 463–474.
[17] Koike, H., Ikeuchi, M., Hiyama, T. and Inoue, Y. (1989) FEBS Lett. 253, 257–263.

[18] Hoffman, N.E., Pichersky, E., Malik, V.S., Castresana, C., Ko, K., Darr, S.C. and Cashmore, A.R. (1987) Proc. Natl. Acad. Sci. USA 84, 8844–8848.
[19] Stayton, M.M., Brosio, P. and Dunsmuir, P. (1987) Plant Mol. Biol. 10, 127–137.
[20] Pichersky, E., Brock, T.G., Nguyen, D., Hoffman, N.E., Piechulla, B., Tanksley, S.D. and Green, B.R. (1989) Plant Mol. Biol. 12, 257–270.
[21] Scheller, H.V., Svendsen, I. and Møller, B.L. (1989) J. Biol. Chem. 264, 6929–6934.
[22] Mullet, J.E., Burke, J.J. and Arntzen, C.J. (1980) Plant Physiol. 65, 814–822.
[23] Ikeuchi, M. and Inoue, Y. (1988) Plant Cell Physiol. 29, 1233–1239.
[24] Ikeuchi, M., Takio, K. and Inoue, Y. (1989) FEBS Lett. 242, 263–269.
[25] Ikeuchi, M. and Inoue, Y. (1989) FEBS Lett. 251, 155–160.
[26] Ikeuchi, M. and Inoue, Y., in preparation.
[27] Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubaybashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. and Sugiura, M. (1986) EMBO J. 5, 2043–2049.
[28] Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.R., Meng, B.Y., Li, Y.Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K. and Sugiura, M. (1989) Mol. Gen. Genet. 217, 185–194.
[29] Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. (1986) Nature 322, 572–574.
[30] Koike, H., Ikeuchi, M., Hiyama, T., Mamada, K. and Inoue, Y. (1990) in: Current Research in Photosynthesis (Baltscheffsky, M. ed.) in press.
[31] Sasaki, Y., Nagano, Y., Morioka, S., Ishikawa, H. and Matsuno, R. (1989) Nucleic Acids Res. 17, 6217–6227.
[32] Scheller, H.V., Okkels, J.S., Høj, P.B., Svendsen, I., Roepstorff, P. and Møller, B.L. (1989) J. Biol. Chem. 264, 18402–18406.
[33] Golbeck, J.H. and Cornelius, J.M. (1986) Biochim. Biophys. Acta 849, 25–31.
[34] Hoshina, S., Sue, S., Kunishima, N., Kamide, K., Wada, K. and Itoh, S. (1989) FEBS Lett. 258, 305–308.
[35] Franzén, L.-G., Frank, G., Zuber, H. and Rochaix, J.-D. (1989) Mol. Gen. Genet., in press.
[36] Girard, J., Chua, N.H., Bennoun, P., Schmidt, G. and Delosme, M. (1980) Curr. Genet. 2, 215–221.