

Short Sequence-Paper

Classification of rice allergenic protein cDNAs belonging to the α -amylase/trypsin inhibitor gene family [☆]

Angelina M. Alvarez, Takahiro Adachi ^{*}, Masayuki Nakase, Naohito Aoki, Ryo Nakamura, Tsukasa Matsuda

Department of Applied Biological Sciences, School of Agricultural Sciences, Nagoya University, Chikusa, Nagoya 464-01, Japan

Received 23 February 1995; revised 10 April 1995; accepted 13 April 1995

Abstract

Seven cDNA clones encoding rice allergenic proteins were newly isolated. Comparison of the sequences of ten cDNA clones, including the previously isolated three clones results in their classification into four subfamilies. Homologies in the nucleotide sequences among and within subfamilies are 70–85% and above 95%, respectively. A sequence of twenty five amino-acid residues at the C-terminal proximal region is highly conserved among all clones and resembles that of plant lipid transfer proteins.

Keywords: Rice allergen; Multigene family; α -Amylase/trypsin inhibitor family; Plant lipid transfer protein

Rice allergens (RAs) have been isolated as antigens to serum IgE antibodies of patients hypersensitive to ingested rice [1,2]. These antigens in rice endosperm are salt-soluble proteins with microheterogeneity in molecular size (14–16 kDa) and isoelectric point (pI 6–8) [3]. Three cDNA and two genomic clones encoding RAs [4,5] have been isolated, suggesting that RAs are encoded by a multigene family and that the protein microheterogeneity may not be due to mere post-translational modification. RAs exhibit sequence homology with wheat α -amylase and barley trypsin inhibitors indicating that RAs are members of the α -amylase/trypsin inhibitor family [5,6]. It is still uncertain whether RAs inhibit α -amylases or trypsin, although a rice salt-soluble protein of about 14 kDa has recently been reported to inhibit α -amylases [7].

Because at least seven immunologically cross-reactive RAs were detected in rice seed proteins [3], the present study required a further search for members of the RA multigene family in a cDNA library of maturing rice seeds

(*O. sativa* L. var *japonica* cv *Nipponbare*) [5]. Therefore, the library was screened with a 370 bp ³²P-labeled *SalI*-*EcoRI* fragment of a formerly identified cDNA, RA17. The inserts of 450–600 bp from positive clones were subcloned in pUC119 and sequenced using the dideoxy-chain termination method as described previously [4,5,8]. Seven novel RA cDNA clones were obtained: three full-length, referred to here as RA5B, RA14B and RA16, and four clones with truncations in the 5' end ('5'-truncated' clones), namely RA14C, RA14d, RA14e and RA14f (Fig. 1 and Fig. 2). Several other clones were also isolated, but they were identical to the seven mentioned.

Table 1 summarizes the sequence identities among seven clones, i.e., six full-length (three previous clones, RA5, RA14 and RA17, and three new clones, RA5B, RA14B and RA16) and a new but 5'-truncated clone, RA14C. Although RA14C is truncated, it contains the whole region coding for a mature protein. Thus it is included in the seven.

The nucleotide sequences of RA14B and the 5'-truncated clones are more than 95% identical to RA14 while the sequence of RA5B exhibits the same extent of identity to that of RA5. Apparently, these clones constitute two RA subfamilies that may be represented by RA14 and RA5. Among the RA5 subfamily, the RA14 subfamily, RA16 and RA17, the identities yield only 70–85%. RA16 and RA17 might be representatives also of two different sub-

[☆] The nucleotide sequence data reported in this paper will appear in the GSDS, DDBJ, EMBL and NCBI nucleotide sequence Data Libraries under the following accession numbers: D42139 (RA14B), D42140 (RA14C), D42141 (RA5B) and D42142 (RA16), D43657 (RA14d), D43658 (RA14e) and D43659 (RA14f).

^{*} Corresponding author. E-mail: i4523a@nucc.cc.nagoya-u.ac.jp. Fax: +81 52 7894120.

[illegible]

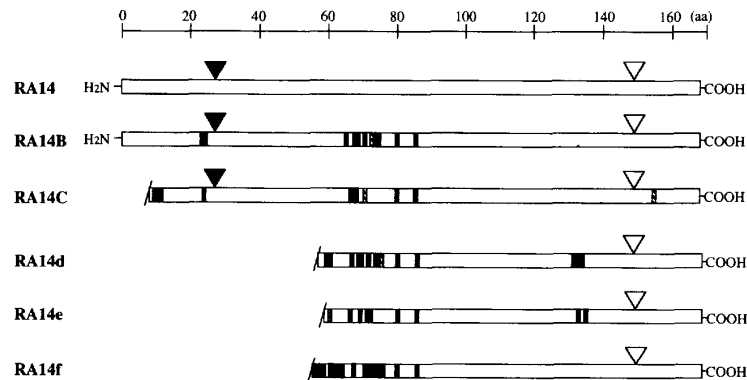


Fig. 2. Location of differences between the deduced amino acids of RA14 and members of RA14 subfamily. White boxed area represents the sequence of RA14 (reference). Deviations from the amino acids of RA14 are marked with hatched and solid boxes to denote insertion and replacement, respectively. Solid triangles point to the putative signal peptide cleavage site, while open triangles indicate putative *N*-glycosylation site (-Asn-Gly-Thr-). Sequence are aligned to keep maximum identity.

families, even though other members of these potential subfamilies have not yet been obtained. A comparison of all sequences shows that the RA14 subfamily comprises the majority in the RA multigene family. A schematic representation (Fig. 2) of the deduced amino-acid sequences of the members of the RA14 subfamily displays that changes are concentrated around the middle region of the polypeptides; however, the C-terminal side is almost conserved. A putative *N*-glycosylation site, Asn-Gly-Thr, exists near the C-terminus of all members of the RA14 subfamily, whereas in the other clones, Thr is replaced by Pro, Gly or Val.

Alignment of the sequences of RA5B, RA14B, RA14C and RA16 to the sequence of the original clone, RA17, manifests that these clones diverge from RA17 due to various replacements and insertions/deletions occurring randomly in the sequences except for a 70 bp region near the 3' end that is highly conserved among the five clones (Fig. 1). Because this is the only region with a conserved motif consisting of 25 amino-acid residues and because the Chou-Fasman analysis [9] predicts a helical structure in this region, we used a computer search of a protein sequence database, SWISSPROT, to look for similar sequences in other proteins. The search yielded the informa-

Table 1
Percentage identities of RA cDNA clones at the nucleotide (DNA) and deduced amino-acid (AA) levels of the mature protein-encoding region

Isolated cDNA	Level	RA14 ^a	RA14B	RA14C	RA5 ^a	RA5B	RA16 ^a	RA17 ^a
RA14B	DNA	96.7	— ^b	97.3	81.2	81.3	75.4	83.0
	AA	91.4		94.3	76.5	73.6	69.5	76.4
RA14C	DNA	97.1	D	— ^b	81.1	79.8	76.6	77.1
	AA	95.0			75.5	76.6	70.9	77.1
RA5B	DNA	81.0	D	D	95.1	— ^b	83.0	77.9
	AA	72.7			93.3		77.6	69.3
RA16	DNA	75.6	D	D	82.1		— ^b	73.5
	AA	68.7			71.8	D		67.7

^a cDNA clones previously isolated [1,5].

^b Same cDNA clones are compared.

D comparison is done in another column.

Fig. 1. Comparison of the nucleotide sequences and deduced amino-acid sequences of new RA cDNA clones with reference to RA17. Identical nucleotides are marked with dots while identical amino acids are marked by asterisks. Arrowhead points to the putative signal peptide cleavage site. The initiation and termination codons, and the putative *N*-glycosylation site are boxed. The conserved cysteine residues are shown with double underlines. The polyadenylation signal is underlined. The conserved domain resembling plant lipid-transfer proteins is marked with an overhead line. The estimated molecular weights (M_r) of the clones are indicated most right of the termination codon.

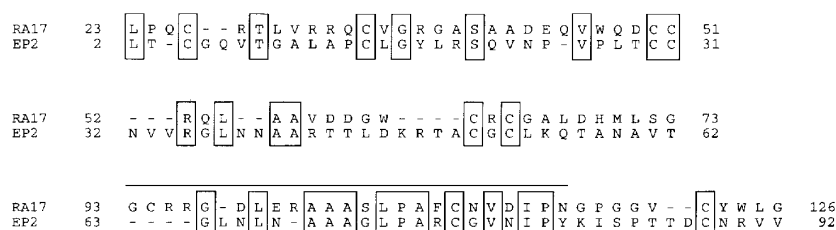


Fig. 3. Alignment of the deduced amino acids of RA17 and a lipid transfer protein in carrot (EP2). The conserved region among RA cDNAs is marked with an overhead line. Identical residues (including conserved cysteines) are boxed. Dashes are introduced for maximum alignment. Amino-acid numbers are assigned based on the (putative) mature protein-encoding sequences.

tion that the homologous sequence of 19 amino-acid residues is present in plant lipid-transfer proteins. Nine to twelve out of nineteen (42–63%) amino-acid residues in different plant lipid-transfer proteins found in maize [10–12], spinach [13], castor bean [14,15], barley [16], carrot [17] and rice [18] are identical to RAs, in which the best match is the sequence of a carrot lipid-transfer protein, EP2 (Fig. 3). Comparison of the sequences also reveals that the positioning of cysteine residues in EP2 as well as in other plant lipid-transfer proteins is quite reminiscent of RAs and α -amylase/trypsin inhibitors (Fig. 3). This suggests that the disulfide bonds might be located in similar positions in lipid-transfer proteins and α -amylase/trypsin inhibitor family and that they may be grouped into one superfamily. However, the molecular mass of plant lipid transfer proteins (9 to 10 kDa) is lower than those deduced from rice allergen cDNAs.

A role of plant lipid transfer proteins has been proposed in the transport of cutin monomers through the extracellular matrix to sites of cutin synthesis on the outer surface of epidermal cells [17]. Plant lipid-transfer proteins have been reported to bind various kinds of fatty acids and lipids [19–22]. They also seem to possess anti-bacterial and anti-fungal activities in vitro [23,24]. No common or similar functions between lipid transfer proteins and RAs or α -amylase/trypsin inhibitors have ever before been reported. Further studies on RA and EP2/plant lipid transfer protein functions would be needed to demonstrate the biological implication of the conserved region in these proteins.

A.M. Alvarez extends her gratitude to the Ministry of Education, Science and Culture of Japan for the financial assistance during the conduct of this work.

References

- [1] Matsuda, T., Sugiyama, M., Nakamura, R. and Torii, S. (1988) *Agric. Biol. Chem.* 52, 1465–1470.
- [2] Shibasaki, M., Suzuki, S., Nemoto, H. and Kuroume, T. (1979) *J. Allergy Clin. Immunol.* 64, 259–265.
- [3] Matsuda, T., Nomura, R., Sugiyama, M. and Nakamura, R. (1991) *Agric. Biol. Chem.* 55, 509–513.
- [4] Adachi, T., Izumi, H., Yamada, T., Tanaka, K., Takeuchi, S., Nakamura, R. and Matsuda, T. (1993) *Plant Mol. Biol.* 21, 239–248.
- [5] Izumi, H., Adachi, T., Fujii, N., Matsuda, T., Nakamura, R., Tanaka, K., Urisu, A. and Kurosawa, Y. (1992) *FEBS Lett.* 302, 213–216.
- [6] Adachi, T., Matsuda, T., Nakamura, R., Alvarez, A.M. and Garcia, V.V. (1994) in *Postharvest Biochemistry of Plant Food Materials in the Tropics* (Uritani, I., Garcia, V.V., Mendoza, E.M.T., eds), pp. 3–15, Japan Scientific Societies Press, Tokyo.
- [7] Feng, G.H., Chen, M., Kramer, K.J. and Reeck, G.R. (1991) *Cereal Chem.* 68, 516–521.
- [8] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- [9] Chou, P.Y. and Fasman, G. (1978) *Adv. Enzymol.* 47, 145–147.
- [10] Tchang, F., This, P., Stiefel, V., Arondel, V., Morch, M.D., Pages, M., Puigdomenech, P., Grellet, F., Delseny, M., Bouillon, P., Huet, J.C., Guerbet, F., Beauvais-Cante, F., Duranton, H., Pernollet, J.C. and Kader, J.-C. (1988) *J. Biol. Chem.* 263, 16849–16855.
- [11] Sossountzov, L., Avila, L.R., Vignols, F., Jolliot, A., Arondel, V., Tchang, F., Grosbois, M., Guerbet, F., Miginiac, E., Delseny, M., Puigdomenech, P. and Kader, J.-C. (1991) *Plant Cell* 3, 923–933.
- [12] Arondel, V., Tchang, F., Baillet, B., Vignols, F., Grellet, F., Delseny, M., Kader, J.-C. and Puigdomenech, P. (1991) *Gene* 99, 133–136.
- [13] Bernhard, W.R., Thoma, S., Botella, J. and Somerville, C.R. (1991) *Plant Physiol.* 95, 164–170.
- [14] Tsuboi, S., Suga, T., Takishima, K., Mamiya, G., Ozeki, Y. and Yamada, M. (1991) *J. Biochem.* 110, 823–831.
- [15] Takishima, K., Watanabe, S., Yamada, M., Suga, T. and Mamiya, G. (1988) *Eur. J. Biochem.* 177, 241–249.
- [16] Mundy, J. and Rogers, J.C. (1986) *Planta* 169, 51–63.
- [17] Sterk, P., Boonij, H., Schellekens, G.A., Van Kammen, A., De Vries, S.C. (1991) *The Plant Cell* 3, 907–921.
- [18] Vignols, F., Lund, G., Pammi, S., Tremousaygue, D., Grellet, F., Kader, J.-C., Puigdomenech, P. and Delseny, M. (1994) *Gene* 142, 265–270.
- [19] Kader, J.C. (1990) *Subcell. Biochem.* 16, 69–111.
- [20] Nishida, I. and Yamada, M. (1986) *Biochim. Biophys. Acta* 813, 298–306.
- [21] Rickers, J., Spener, F. and Kader, J.C. (1985) *FEBS Lett.* 180, 29–32.
- [22] Rickers, J., Tober, I. and Spener, F. (1984) *Biochim. Biophys. Acta* 784, 313–319.
- [23] Molina, A., Segura, A. and Garcia-Olmedo, F. (1993) *FEBS Lett.* 316, 119–122.
- [24] Terras, F.R.G., Goderis, L.J., Van Leuven, F., Vanderleyden, J., Cammue, B.P.A. and Broekaert, W.F. (1992) *Plant Physiol.* 100, 1055–1058.