Volume 302, number 3, 213–216 FEBS 11026 May 1992

Nucleotide sequence of a cDNA clone encoding a major allergenic protein in rice seeds

Homology of the deduced amino acid sequence with members of α-amylase/trypsin inhibitor family

Hidehiko Izumi[a], Takahiro Adachi[a], Noboru Fujii[a,*], Tsukasa Matsuda[a], Ryo Nakamura[a], Kunisuke Tanaka[b], Atsuo Urisu[c] and Yoshikazu Kurosawa[d]

[a]*Department of Food Science and Technology, School of Agriculture, Nagoya University, Nagoya, Aichi 464-01, Japan,* [b]*Department of Biochemistry, College of Agricultural Chemistry, Kyoto Prefectural University, Kyoto, Kyoto 606, Japan* and [c]*Department of Pediatrics and* [d]*Institute for Comprehensive Medical Science, Fujita-Gakuen Health University, Toyoake, Aichi 470-11, Japan*

Received 12 February 1992; revised version received 30 March 1992

A cDNA clone of rice major allergenic protein (RAP) was isolated from a cDNA library of maturing rice seeds. The cDNA had an open reading frame (486 nucleotides) which coded a 162 amino acid residue polypeptide comprising a 27-residue signal peptide and a 135-residue mature protein of M_r 14,764. The deduced amino acid sequence of RAP showed a considerable similarity to barley trypsin inhibitor [1983, J. Biol. Chem. 258, 7998–8003] and wheat α-amylase inhibitor [1981, Phytochemistry 20, 1781–1784].

Rice allergen; Seed protein; cDNA cloning; α-Amylase-inhibitor family; Sequence homology

1. INTRODUCTION

Several studies on allergic reaction to cereal grains have suggested that sera from patients with atopic food allergy contain immunoglobulin E (IgE) reacting with rice proteins, and that salt-soluble proteins, albumins and globulins have a high degree of allergenic activity [1–4].

The authors have isolated a rice seed protein of about 16 kDa with reactivity for IgE in several rice-allergic patients [5]. Antibodies raised against the 16K protein detected several immunologically cross-reactive proteins in the rice soluble protein fraction [6]. The rice 16K protein was later identified as a major allergen in rice seeds based on the results of radio allergosorbent test (RAST), RAST inhibition and immunoblotting analyses using 31 patients' sera [7,8].

In this study, a cDNA encoding the rice major allergenic protein (RAP) was isolated and its nucleotide sequence was determined. It revealed considerable similarity of the deduced amino acid sequence of RAP to the sequences of member proteins of α-amylase/trypsin inhibitor family [9] in cereal and legume seeds. This is the first report on cloning and sequencing of an allergenic protein from α-amylase/trypsin inhibitor family.

2. MATERIALS AND METHODS

RNA was extracted from maturing rice seeds harvested at 20 days after flowering, and Poly(A^+) RNA was purified as described in [10]. A λgt11 cDNA library was constructed according to the published procedure in [11]. The cDNA library was screened by using a ^{32}P-labeled synthetic oligonucleotide mixture, 5'-NNRTGNACYTGRT-GRTGRTC-3' (R, Y and N represent purine, pyrimidine, and purine or pyrimidine nucleotides, respectively), which contained the complementary nucleotide sequence to the N-terminal amino acid sequence of RAP, Asp^1-His-His-Gln-Val-Tyr-Ser^7 (Fig. 1). RAP was isolated as described in [5]. Thereafter, the N-terminal amino acid sequence of the purified RAP was determined by an ABI protein sequencer, type 477-120A (Applied Biosystems). cDNA inserts were subcloned into pUC118 or 119, and sequenced by the dideoxy chain-termination method [12].

3. RESULTS AND DISCUSSION

3.1. *Cloning and sequencing of RAP cDNA*

A cDNA library from maturing rice endosperm mRNA was screened with the synthetic 20-mer. Hybridization of the ^{32}P-labeled probes led to the identification of several putative RAP clones. Size analysis of the

Present address: Tokyo Research Institute, Kyowa Hakko Kogyo Co. Ltd., Asahi-cho 3-6-6, Machida, Tokyo 194, Japan.

Correspondence address: T. Matsuda, Department of Food Science and Technology, School of Agriculture, Nagoya University, Nagoya, Aichi 464-01, Japan. Fax: (81) (52) 782 9162.

```
                                                    TATTTTTCACAAACTAAGCAAAATATGGCTTCCAACAAGGTAGTG
                                                                          M  A  S  N  K  V  V
        .        .        .        .        .        .        .        .        .        .
  46    TTCTCGGTATTGCTCCTCGTCGTTCTCTCCGTGCTCGCGGCGGCGATGGCGACCATGGCG
 -20     F  S  V  L  L  L  V  V  L  S  V  L  A  A  A  M  A  T  M  A
        .        .        .        .        .        .        .        .        .        .
 106    GACCACCACCAAGTCTACAGCCCCGGTGAGCAATGTCGGCCAGGAATAAGCTATCCGACA
   1     D  H  H  Q  V  Y  S  P  G  E  Q  C  R  P  G  I  S  Y  P  T
        .        .        .        .        .        .        .        .        .        .
 166    TACTCACTCCCGCAATGTCGAACATTGGTGAGGCGCCAATGCGTTGGCCGTGGCGCCAGC
  21    Y  S  L  P  Q  C  R  T  L  V  R  R  Q  C  V  G  R  G  A  S
        .        .        .        .        .        .        .        .        .        .
 226    GCCGCGGACGAGCAAGTCTGGCAAGACTGCTGCCGGCAGCTCGCCGCCGTCGACGACGGC
  41    A  A  D  E  Q  V  W  Q  D  C  C  R  Q  L  A  A  V  D  D  G
        .        .        .        .        .        .        .        .        .        .
 286    TGGTGCAGGTGCGGGGCGCTCGATCACATGCTGTCAGGCATCTACAGGGAGCTCGGCGCC
  61    W  C  R  C  G  A  L  D  H  M  L  S  G  I  Y  R  E  L  G  A
        .        .        .        .        .        .        .        .        .        .
 346    ACCGAGGCCGGGCACCCCATGGCCGAGGTGTTCCCCGGCTGCCGGAGAGGGGACTTGGAG
  81    T  E  A  G  H  P  M  A  E  V  F  P  G  C  R  R  G  D  L  E
        .        .        .        .        .        .        .        .        .        .
 406    CGCGCGGCGGCGAGCCTCCCGGCGTTCTGCAACGTGGACATCCCCAATGGGCCTGGTGGT
 101    R  A  A  A  S  L  P  A  F  C  N  V  D  I  P  N  G  P  G  G
        .        .        .        .        .        .        .        .        .        .
 466    GTCTGCTACTGGCTTGGGTATCCTAGGACCCCAAGAACCGGTCACTAGGCTACTATAGTT
 121    V  C  Y  W  L  G  Y  P  R  T  P  R  T  G  H  #
        .        .        .        .        .        .        .        .        .        .
 526    AGCTGTGTGTGTGTGACTATGTGGGGTTGCTAAATAATTAGTGCTTTCTTTTGTCAGGAA
        .        .        .        .        .        .        .        .        .        .
 586    GCATCTATTTATATATATGTGGTGAATAAATGATGAACTTCAATGTTCTTCTG
```

Fig. 1. Nucleotide and deduced amino acid sequences of a cDNA of a rice major allergenic protein (RAP). Numbers in the left margin refer to arbitrary positions of the amino acid and nucleotide, respectively. A possible ATG start codon and a putative polyadenylation signal are underlined, and a TAG stop codon at the end of the open reading frame is indicated by the symbol #. The deduced amino acid sequence is shown by a one-letter code under the corresponding codon.

cDNA inserts of these clones indicated that one of them was about 1,200 nucleotides in length and the others about 300 nucleotides or less. The nucleotide sequence showed that the long cDNA insert consisted of a long 5′ untranslated region followed by an open reading frame of RAP. The nucleotide sequence covering the open reading frame and the deduced amino acid sequence are shown in Fig. 1. The open reading frame consisting of 486 nucleotides encoded 162 amino acids, including a putative signal peptide of 27 amino acids and a 135 residue mature protein of M_r 14,764 in agreement with the estimate of M_r from SDS/PAGE [5]. The amino acid sequence deduced from nucleotide 106 to 165 corresponded exactly with that determined directly for the 20 N-terminal amino acids of the isolated RAP. Cleavage of the precursor at that position between Ala^{-1} and Asp^{1} follows Von Heijne's rule for the cleavage site of a signal sequence [13]. The putative signal peptide of 27 residues had properties of typical signal peptides; a basic amino acid, Lys^{-23}, followed by hydrophobic amino acids.

3.2. *Comparison of the amino acid sequences of RAP and members of α-amylase/trypsin inhibitor family*

A search of the protein sequence database for NBRF, EMBL and GenBank revealed sequence similarities between RAP and eight members of α-amylase/trypsin inhibitor family, namely the monomeric inhibitors of heterologous α-amylase from wheat [14–16], the trypsin inhibitors from barley [17] and maize [18], α-amylase/trypsin inhibitors from rye [19] and ragi [20], and the glutamine-rich storage protein from castor bean [21]. In Fig. 2, the deduced amino acid sequence of RAP is compared with the sequences of three members of the family. The sequence identities between RAP and wheat α-amylase inhibitor and between RAP and barley trypsin inhibitor are about 40% and 20%, respectively. The RAP sequence aligned well with wheat α-amylase inhib-

```
                    .     10    .     20    .     30    .     40    .
RAP                 DHHQVYSPGEQCRPGISYPTYSLPQCRTLVRRQ.CVGRGASAADEQVW
Wheat AI 28             SGPWSWCNPATGYKVSALTGCRAMVKLQ.CVGSQVP...EAVL
Barley TI                 FGDSCAPGDALPHNPLRACRTYVVSQICHQGPRL.LTSDMK
Castor Bean             PSQQGCRGQIQ.EQQNLRQCQEYIKQQVSGQGPRR  QERSL

                   50     .    60     .    70     .    80     .    90    .
RAP                 QDCCRQLAAVDDGWCRCGALDHMLSGIYRELGATEAGHPMAEVFPGCR
Wheat AI 28         RDCCQQLADINNEWCRCGDLSSMLRAVYQELGVRE.GK...EVLPGCR
Barley TI           RRCCDELSAIP.AYCRCEALRIIMQGVVTWQGAFE.GAYF.KDSPNCP
Castor Bean         RGCCDHLKQMQ.SQCRCEGLRQAIQQQQL.QGQNVFEAF.........

                        100     .     110     .     120     .     130    .
RAP                 RGDLERAAASL..PAFCNVDIPNGPG...GVCY..WLGYPRTPRTGH
Wheat AI 28         KEVMKLTAASV..PEVCKVPIPNPSGDRAGVCYGDWCAYPDV
Barley TI           RERQTSYAANLVTPQECNLGTIH..G.SA.YCPELQPGYG
Castor Bean         .....RTAANL..PSMC........GVSPTQCRF
```

Fig. 2. Alignment of the deduced amino acid sequence of RAP with sequences of α-amylase/trypsin inhibitor family proteins, namely, wheat α-amylase inhibitor 0.28 (AI 28) [14], barley trypsin inhibitor (TI) [17] and castor bean glutamine-rich storage protein [21]. The sequences are aligned for maximum homology, resulting in several gaps (shown as a dot) which may represent insertions/deletions. Numbering refers the sequence for RAP. Conserved cysteine residues are shown in boldface. The castor-bean protein consists of two polypeptide chains of 34 and 61 residues.

itor and barley trypsin inhibitor with 10 cystein residues being conserved. These results suggest the possible divergent evolution of RAP and the α-amylase/trypsin inhibitors from a common ancestor. Thus, while the sequence comparisons clearly show that RAP is related to the α-amylase/trypsin inhibitor family of proteins, its relationship to wheat α-amylase inhibitor is closer than that to barley trypsin inhibitor and that observed between wheat α-amylase inhibitor and barley trypsin inhibitor, which show about 20% identity.

3.3. *α-Amylase/trypsin inhibitor family and major seed allergens*

Members of the α-amylase inhibitor family from wheat endosperm have recently been reported to be major allergens associated with baker's asthma [22], and a barley salt-soluble protein of 14.5 kDa, which inhibits insect α-amylase has also been identified as a major antigen recognized by IgE of sera from baker's asthma patients [23]. Furthermore, the castor bean storage protein reported to be allergenic [24] is a member of the α-amylase-trypsin inhibitor family [21,25]. In the present study, the major allergenic protein in rice seeds was found to be a member of this inhibitor family. Thus, several member proteins from four different species, namely wheat, barley, rice and castor bean have been identified as allergenic components, suggesting that the α-amylase/trypsin inhibitor family proteins are potentially prominent allergens in cereal and legume seeds.

REFERENCES

[1] Hoffman, D.R. (1975) Immunochemistry 12, 535–538.
[2] Shibasaki, M., Suzuki, S., Nemoto, H. and Kuroume, T. (1979) J. Allergy Clin. Immunol. 64, 259–265.
[3] Watanabe, M., Miyakawa, J., Ikezawa, Z., Suzuki, Y., Hirano, T., Yoshizawa, T. and Arai, S. (1990) J. Food Sci. 55, 781–783.
[4] Limas, G.G., Salinas, M., Moneo, I., Fischer, S., Wittmann-Liebold and Mendez, E. (1990) Planta 181, 1–9.
[5] Matsuda, T., Sugiyama, M., Nakamura, R. and Torii, S. (1988) Agric. Biol. Chem. 52, 1465–1470.
[6] Matsuda, T., Nomura, R., Sugiyama, M. and Nakamura, R. (1991) Agric. Biol. Chem. 55, 509–513.
[7] Urisu, A., Wada, E., Kondo, Y., Horiba, F., Tsuruta, M., Yasaki, T., Yamada, K., Masuda, S., Komada, H., Yamada, M., Torii, S. and Nakamura, R. (1991) Jpn. J. Allergol. 40, 1370–1376.
[8] Urisu, A., Yamada, K., Masuda, S., Komada, H., Wada, E., Kondo, Y., Horiba, F., Yazaki, T., Yamada, H., Torii, S. and Nakamura, R., Int. Arch. Allergy Appl. Immunol. (in press).
[9] Barber, D., Sanchez-Monge, R., Garcia-Olmedo, F., Sakedo, G. and Mendez, E. (1986) Biochim. Biophys. Acta 873, 147–151.
[10] Yamagata, H., Tamura, K., Tanaka, K. and Kasai, Z. (1986) Plant Cell Physiol. 27, 1419–1422.
[11] Young, R.A. and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194–1198.
[12] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[13] Von Heijne, G. (1983) Eur. J. Biochem. 133, 17–21.
[14] Kashlan, N. and Richardson, M. (1981) Phytochemistry 20, 1781–1784.
[15] Maeda, K., Hase, T. and Matsubara, H. (1983) Biochim. Biophys. Acta 743, 52–57.
[16] Maeda, K., Kababayashi, S. and Matsubara, H. (1985) Biochim. Biophys. Acta 828, 213–221.
[17] Odani, S., Koide, T. and Ono, T. (1983) J. Biol. Chem. 258, 7998–8003.

[18] Mahoney, W.C., Hermodson, M.A., Jones, B., Powers, D.D., Corfman, R.S. and Reeck, G.R. (1984) J. Biol. Chem. 259, 8412–8416.
[19] Lyos, A., Richardson, M., Tatham, A.S. and Shewry, P.R. (1987) Biochim. Biophys. Acta 915, 305–313.
[20] Campos, F.D.A. and Richardson, M. (1983) FEBS Lett. 152, 300–304.
[21] Sharief, F.S. and Li, S.S.-L. (1982) J. Biol. Chem. 257, 14753–14759.
[22] Gomez, L., Martin, E., Hernandez, D., Sanchez-Monge, R., Barber, D., Pozo, V., Andres, B., Armentia, A., Lahoz, C., Salcedo, G. and Palomino, P. (1990) FEBS Lett. 261, 85–88.
[23] Barber, D., Sanchez-Monge, R., Gomez, L., Carpizo, J., Armentia, A., Lopez-Otin, C., Juan, F. and Salcedo, G. (1989) FEBS Lett. 248, 119–122.
[24] Spies, J.R. (1974) J. Agric. Food Chem. 22, 30–36.
[25] Li, S.S.-L., Lin, T.T.-S. and Forde, M.D. (1977) Biochim. Biophys. Acta 492, 364–369.