Primary structures of two ribonucleases from ginseng calluses

New members of the PR-10 family of intracellular pathogenesis-related plant proteins[1]

Gennady P. Moiseyev[a], Larisa I. Fedoreyeva[b], Yuri N. Zhuravlev[c], Elena Yasnetskaya[c], Peter A. Jekel[d], Jaap J. Beintema[d,*]

[a] *Engelhardt Institute of Molecular Biology, Academy of Sciences, Vavilov Street 32, 117984 Moscow B-334, Russia*
[b] *Pacific Institute of Bioorganic Chemistry, Russian Academy of Sciences, 690022 Vladivostok-22, Russia*
[c] *Institute of Biology and Soil Science of the Far East Department of the Russian Academy of Sciences, 690022 Vladivostok-22, Russia*
[d] *Department of Biochemistry, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands*

Received 12 February 1997; revised version received 14 March 1997

Abstract **The amino acid sequences of two ribonucleases from a callus cell culture of *Panax ginseng* were determined. The two sequences differ at 26% of the amino acid positions. Homology was found with a large family of intracellular pathogenesis-related proteins, food allergens and tree pollen allergens from both dicotyledonous and monocotyledonous plant species. There is about 30% sequence difference with proteins from species belonging to the same plant order (*Apiales*: parsley and celery), 60% with those from four other dicotyledonous plant orders and about 70% from that of the monocotyledonous asparagus. More thorough evolutionary analyses of sequences lead to the conclusion that the general biological function of members of this protein family may be closely related to the ability to cleave intracellular RNA and that they have an important role in cell metabolism. As the three-dimensional structure of one of the members of this protein family has been determined recently [Gajhede et al., Nature Struct Biol 3 (1996) 1040–1045], it may be possible to assign active-site residues in the enzyme molecule and make hypotheses about its mode of action. Structural features in addition to the cellular site of biosynthesis indicate that this family of ribonucleases is very different from previously investigated ones.**

Key words: Allergen; Amino acid sequence; Ginseng; Pathogenesis-related protein; Ribonuclease

1. Introduction

One of the biological functions of ribonucleases is to digest extracellular RNA. However, several studies [1,2] indicate that these enzymes also play important intracellular roles, e.g. as cytotoxins in cell metabolism and proliferation. The majority of investigated ribonucleases are synthesized at the rough endoplasmic reticulum with signal peptide and targeted first to extracellular or equivalent (e.g. lysosomal) destinations, and uptake in the cytoplasm is a secondary process. Few ribonucleases are known and characterized which are made directly in the cytoplasm. One of the few known eukaryotic ones is the 2–5-dependent ribonuclease L [3].

*Corresponding author. Fax: (31) (50) 3634165.

[1] The novel amino acid sequences have been deposited under numbers P80889 and P80890 in the SWISS-PROT protein sequence database.

In our previous paper [4] we described the isolation of a ribonuclease with a molecular mass of about 18 kDa from a callus cell culture of *Panax ginseng* C.A. Mey strain R1. This enzyme cleaves RNA endonucleolytically at the 3′-adjacent phosphodiester bonds of all nucleotides except cytidine, with the formation of nucleoside and oligonucleotide 3′-phosphates via cyclic derivatives as intermediates. Primary structure determination of two fragments of the protein showed high sequence similarity with two intracellular pathogenesis-related (IPR) proteins from parsley [5,6]. These proteins belong to a large homologous family found in both dicotyledonous and monocotyledonous plant species [7], and includes proteins elicited by infection with fungal pathogens or wounding (IPR proteins; now called PR-10 proteins [8]), food allergens from celery and apple [9,10], and tree pollen allergens [11]. These proteins are synthesized in the cytosol without signal peptides [7].

The sequence similarity between ginseng ribonucleases and members of this ubiquitously occurring protein family suggested that its general biological function is also closely related to the capability of cleaving RNA [4]. Spatio-temporal similarities between reported expression patterns of PR-10 genes and ribonuclease genes support this hypothesis [7]. Meanwhile, ribonuclease activity has been demonstrated in birch pollen allergen [12,13].

When we continued our sequence studies on ginseng ribonuclease, we found that our samples contained two homologous components. Here we describe the complete sequences of both components and present an evolutionary analysis of the sequences of members of the IPR protein family, which strongly support the claim that these proteins may be all ribonucleases. Very recently the three-dimensional structure of the major birch allergen Bet v 1 was published [14], which allows drafting hypotheses about a likely location of the active site of the enzyme and its architecture.

2. Materials and methods

Ginseng ribonuclease preparations were isolated as described previously [4].

Sequence studies were performed as described [15,16]. Ribonuclease preparations were used directly or after separation of the two components by reversed-phase HPLC on a C2 column with a linear gradient of 0–60% acetonitrile in 0.1% trifluoroacetic acid. Proteins were cleaved with CNBr in 70% formic acid overnight, or by digestion with trypsin, endoproteinase Glu-C or thermolysin after denaturation

PII S0014-5793(97)00337-2

of the protein by performic acid oxidation. CNBr peptides were cleaved with trypsin or endoproteinase Glu-C. The C-terminal CNBr peptides of the two ribonucleases were isolated by HPLC on a C2 column as described above, and other peptides on a C18 column using the same gradient elution.

Amino acid sequences of peptides were determined by automatic Edman degradation on an Applied Biosystems 477A pulse-liquid sequencer with an on-line 120A phenylthiohydantoin analyzer (Eurosequence BV, Groningen). At several steps during sequence analyses, peptides with N-terminal proline were treated with *ortho*-phthalaldehyde [17] to block contaminants with non-proline N-terminal residues.

Aligned amino acid sequences were analyzed by the neighbor joining method in the CLUSTAL W program package [18] and using maximum parsimony under the heuristic search option (tree bisection-reconnection, local and global swapping, simple and random stepwise addition) using the PAUP computer program, version 3.1.1 [19]. Gaps (deletions) are considered 'missing data'. Searches were done under the 'protpars assumption, in which the number of codon differences between replaced amino acids is taken into account (see also [20]). Bootstrap analyses were used to provide an estimate of tree stability.

3. Results and discussion

Another sequencer run (with treatment with *ortho*-phthalic aldehyde at two proline positions) of the previously investigated ginseng ribonuclease sample [4] extended the number of identified residues to 42 (Fig. 1). However, N-terminal sequence analysis of another ginseng ribonuclease sample showed that it contained two components in about equal quantities. The two components have been named ribonuclease 1 and 2, respectively. CNBr digestion of the mixture also gave a more complex cleavage pattern, as ribonuclease 1 contains only one methionine and yields two fragments (as demonstrated in our previous paper [4]), while ribonuclease 2 has an additional methionine and yields three fragments. The N-terminal CNBr peptide of ribonuclease 1 was isolated by reversed-phase HPLC on a C18 column, and its sequence determination was completed by analysis of several overlapping peptides (Fig. 1). The two N-terminal CNBr peptides of ribonuclease 2 were isolated as a mixture, and their sequences were derived from two sequencer runs on this mixture (with treatment with *ortho*-phthalic aldehyde at two different positions; Fig. 1) and one on the mixture of ribonucleases 1 and 2. Reexamination of the data obtained with the previously investigated ribonuclease sample [4] showed that it was contaminated for about 20% with ribonuclease 2.

The C-terminal CNBr peptides could not be obtained by reversed-phase HPLC on a C18 column. However, two separate sets of overlapping peptides from the region near residues 100–135 could be obtained for both ribonucleases from two enzymic digests of the ribonuclease mixture (Fig. 1).

We could separate ribonucleases 1 and 2 by reversed-phase HPLC on a C2 column, and obtain the C-terminal CNBr peptides of each of them by HPLC on the same column. N-terminal sequence analysis of these two peptides, and of several smaller ones obtained after subdigestion with endopro-

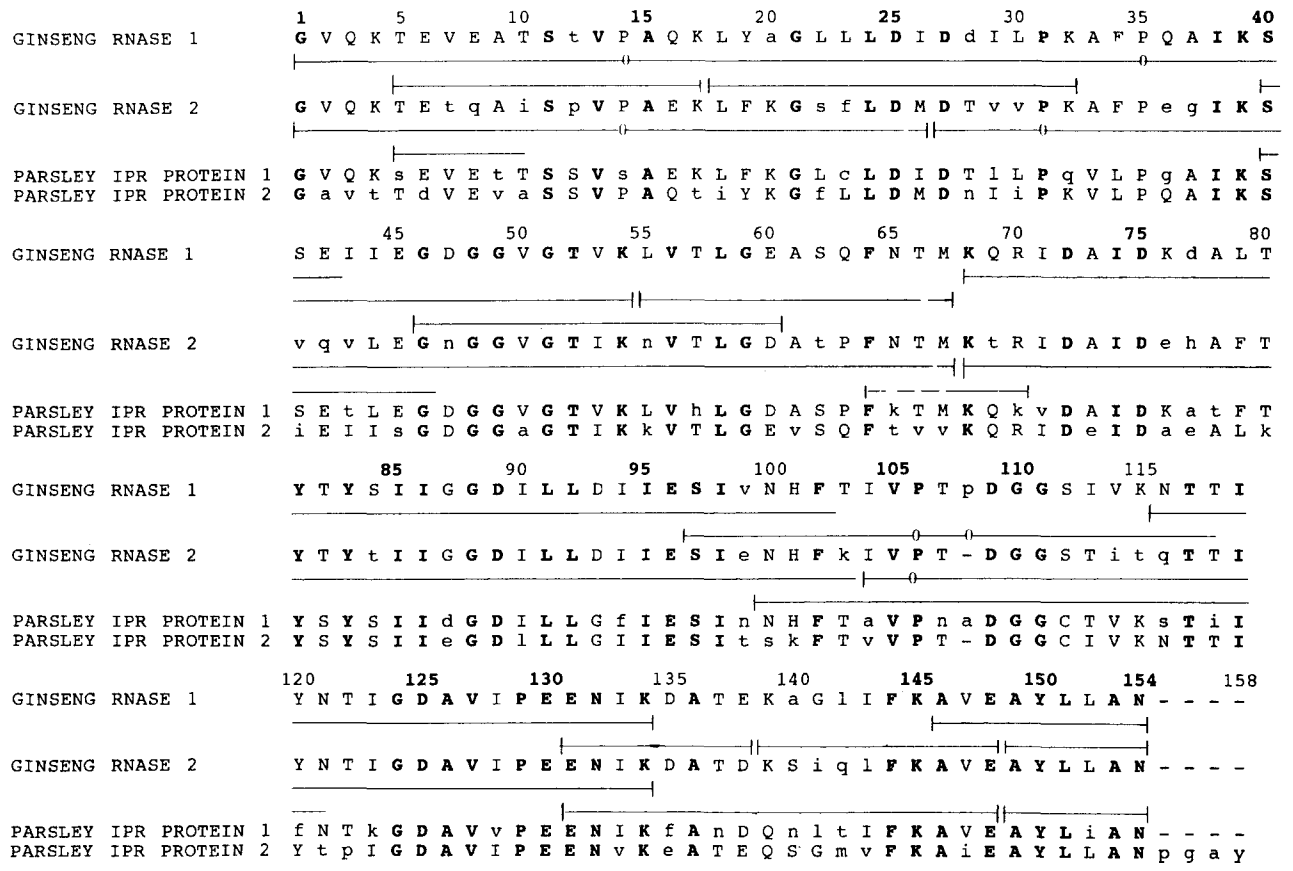

Fig. 1. Amino acid sequences of ginseng ribonucleases 1 and 2 and parsley IPR proteins 1 and 2 [5,6]. Identical residues in the four sequences are indicated in bold capitals. Other identical residues are indicated in capitals, and unique residues in lower case. Completely (1—1) and incompletely (1—) sequenced proteins and peptides are indicated. —○—, proline residues which were treated with *ortho*-phthalaldehyde when located at the N-terminus during sequencing.

teinase Glu-C and peptide isolation by HPLC on a C18 column completed the sequence determinations (Fig. 1). Several overlaps were derived by homology between the two ribonucleases. The overlap between the two C-terminal CNBr peptides of ribonuclease 2 was derived from sequence analysis of a peptide mixture obtained after cleavage of the performic-oxidized protein with thermolysin and reversed-phase HPLC. Although the first three steps of the sequencer run produced a complicated mixture of amino acid derivatives, step 4 clearly showed the presence of methionine sulfone, followed at steps 5–7 by the sequence Lys-Thr-Arg. The proteolytic enzymes used had the expected specificities. The cleavage of the -Pro-Glu-Glu-Asn-(129–132) sequence by endoproteinase Glu-C after the first Glu is in agreement with earlier observations on its specificity [21].

The complete sequences of ginseng ribonucleases 1 and 2 are presented in Fig. 1, aligned with the homologous sequences of parsley intracellular pathogenesis-related proteins 1 and 2 [5,6]. As the two ginseng ribonucleases were separated under denaturing conditions, we have not yet been able to determine if both components are enzymatically active and have similar specific activities. However, as preparations with different ratios of the two components have similar specific activities, we do not expect large differences.

The two ginseng sequences differ at 42, or 26%, of the 154 amino acid residue positions. The difference with the two parsley proteins is about 40%. The sequences have also been aligned with those of other homologous proteins (SWISS-PROT and SWISS-NEW databanks with similarity search using the program FASTA; not shown). Ginseng, parsley and celery (of which a homologous food allergen has been sequenced [9]) belong to the plant order *Apiales*, and their PR-10 proteins differ at 25–40% of the amino acid positions. Homologous proteins from four other dicotyledonous orders (*Fagales*: birch, alder, hazel and hornbeam; *Fabales*: bean, pea, lupin and soybean; *Rosales*: apple; and *Solanales*: potato) differ at about 60%, and that from the monocotyledonous species asparagus (*Liliales*) at about 70% of the positions.

Evolutionary analyses by the neighbor-joining and parsimony methods gave similar results. Generally, different variants in one species (like birch) were monophyletic. Only in the *Fabales* were paralogous relationships observed. However, at the level of plant orders monophyletic groupings were obtained with 98–100% bootstrap values. Fig. 2 shows a most parsimonious tree of PR-10 proteins with the monocotyledonous asparagus sequence [22] as outgroup, the five *Apiales* sequences, and one representative of each of the four other dicotyledonous orders. Bootstrap analysis shows a monophyletic *Apiales* clade, but it is not possible, with these data, to reach unambiguous conclusions about the divergence of the plant orders and about the branching pattern of the five *Apiales* sequences. Both the two parsley and the two ginseng sequences are the products of gene duplications in an ancestral *Apiales* species. However, it cannot be decided whether there has been a gene duplication before the divergence of ginseng and parsley or there have been two independent duplications. Van de Löcht et al. [6] report that there is no sequence similarity between the promoter sequences of the two parsley proteins.

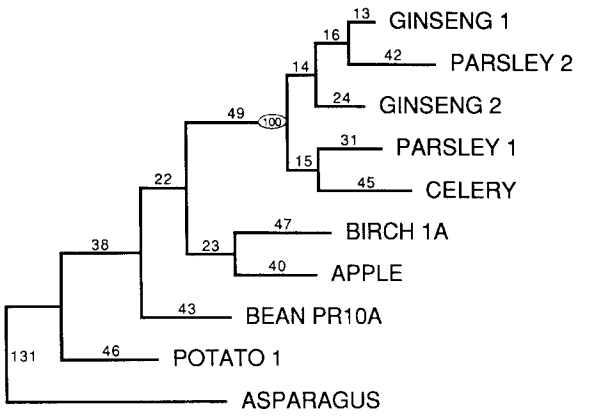

Fig. 2. Most parsimonious tree of PR-10 ribonucleases (five sequences from *Apiales* and one representative from five other plant orders). Two most parsimonious trees were obtained with differing branching patterns in the *Apiales* clade. With a larger number of sequences one most parsimonious tree was obtained, with the branching pattern in the *Apiales* clade as in the presented tree. Numbers above the branches are steps (=amino acid changes which are obtained under the accelerated transformation; ACCTRAN), and the circled number indicates a high bootstrap percentage (>90%).

The orthologous relationships between major lineages in the evolutionary analyses support a common function of the PR-10 proteins. This may be basically an intracellular ribonucleolytic one, with specific additional ones as observed in other ribonuclease superfamilies [1,23]. Few intracellularly synthesized eukaryotic ribonucleases have been identified and characterized yet. Two of the few examples are the 2–5A-dependent human and murine ribonucleases L [3]. The polypeptide chains of these enzymes have chain lengths of about 750 residues, and each includes a ribonuclease domain of about 200 residues, with some sequence similarity with ribonuclease E from *Escherichia coli* but not with the PR-10 plant ribonucleases.

Meanwhile ribonuclease activity has been demonstrated in preparations of birch tree allergens and of the recombinant protein [12,13]. Analysis of pollens collected from individual birch trees has shown that each tree expresses a specific subset of isoforms of pollen allergens [14]. This may reflect a function in recognition and self-incompatibility during the process of fertilization [13]. The allergens will be released from rehydrated pollen in the female reproductive tissue and after uptake in the cytoplasm of the pollen tube they may become cytotoxic ribonucleases [13]. A large family [24] in the superfamily of extracellular T_2 ribonucleases also causes self-incompatibility in plants [23]. These proteins are expressed in the pistils of several plants and are secreted in the style mucilage, where they also could penetrate cells of the pollen tube and act as cytotoxic ribonucleases [23]. Thus these similarities suggest that self-incompatibility in plants may be either female- or male-determined, but in both cases involving cytotoxic proteins which are ribonucleases.

Very recently the X-ray and NMR structure of birch tree Bet v 1 has been published [14]. This allows drafting hypotheses about a likely location of the active site of the enzyme and its architecture. There is no similarity with three-dimensional structures of other ribonucleases, which could already be deduced from the absence of conserved histidine residues in the PR-10 ribonuclease superfamily. Like in the pyrimidine-specific [25] and T_2 ribonuclease [26] superfamilies, few amino acid residues are conserved completely in the PR-10 ribonuclease superfamily. Walter et al. [7] have indicated 12 conserved residues in their alignment of sequences. However,

with additional homologous sequences like the two from ginseng and a few others, this number decreases to nine. Most of them are glycines and a proline, which probably are essential for the polypeptide fold [14]. The only conserved residues with side chains with functional groups and which may be involved in the catalytic reaction are Glu-96, Glu-148 and Tyr-150 (numbering used in Fig. 1). The three-dimensional structure of the birch allergen shows that Glu-148 and Tyr-150 are located at two opposing sides of the long C-terminal helix of the molecule, while the position of Glu-96 is at a rather large distance at the N-terminus of β-strand VI. Gajhede et al. [14], in their description of the three-dimensional structure of Bet v 1, mention the presence of a P-loop structure [27] with the sequence GXGGXGXXK (positions 46–54 in Fig. 1), which may indicate a phosphate binding site and, therefore, a likely place for binding a phosphate group of RNA. The loop is rather well conserved in other members of the PR-10 superfamily. Only in the two potato proteins [28] the loop is shortened by three residues to GDGSIK, and in the asparagus protein [22], the putative phosphate-binding lysine is replaced by an arginine residue.

Future studies will show if ribonuclease activity really is a universal property of PR-10 proteins, and provide information about the catalytic mechanism and other properties of the enzyme.

Acknowledgements: This work was supported by a short-term fellowship to G.P.M. from the Federation of European Biochemical Societies and by the Netherlands Organization for Scientific Research (NWO; Dossier 07-30-162). We thank Dr. K.I. Panov (Moscow) for the donation of a C2 HPLC column, Dr. W.T. Stam (Department of Marine Biology, University of Groningen) and Dr. R.G. Kleineidam (Department of Biochemistry, University of Groningen) for evolutionary analyses of sequences, and Dr. R.N. Campagne (Groningen) for reviewing the manuscript.

References

[1] G. D'Alessio, Trends Cell Biol 3 (1993) 106–109.
[2] R.J. Youle, D. Newton, Y.-N. Wu, M. Gadino, S.M. Rybak, Crit Rev Ther Drug Carrier Syst 10 (1993) 1–28.
[3] A. Zhou, B.A. Hassel, R.H. Silverman, Cell 72 (1993) 753–765.
[4] G.P. Moiseyev, J.J. Beintema, L.I. Fedoreyeva, G.I. Yakovlev, Planta 193 (1994) 470–472.
[5] I.E. Somssich, E. Schmelzer, P. Kawalleck, K. Hahlbrock, Mol Gen Genet 213 (1988) 93–98.
[6] U. Van de Löcht, I. Meier, K. Hahlbrock, I.E. Somssich, EMBO J 9 (1990) 2945–2950.
[7] M.H. Walter, J.-W. Liu, J. Wünn, D. Hess, Eur J Biochem 239 (1996) 281–293.
[8] L.C. Van Loon, W.S. Pierpoint, T. Boller, V. Conejero, Plant Mol Biol Rep 12 (1994) 245–264.
[9] H. Breiteneder, K. Hoffmann-Sommergruber, G. O'Ríordáin, M. Susani, H. Ahorn, C. Ebner, D. Kraft, O. Scheiner, Eur J Biochem 233 (1995) 484–489.
[10] M. Vanek-Krebitz, K. Hoffmann-Sommergruber, M.L.D. Machado, M. Susani, C. Ebner, D. Kraft, O. Scheiner, H. Breiteneder, Biochem Biophys Res Commun 214 (1995) 538–551.
[11] H. Breiteneder, K. Pettenburger, A. Bito, R. Valenta, D. Kraft, H. Rumpold, O. Scheiner, M. Breitenbach, EMBO J 8 (1989) 1935–1938.
[12] A. Bufe, M.D. Spangfort, H. Kahlert, M. Schlaak, W.-M. Becker, Planta 199 (1996) 413–415.
[13] I. Swoboda, K. Hoffmann-Sommergruber, G. O'Ríordáin, O. Scheiner, E. Heberle-Bors, O. Vicente, Physiol Plant 96 (1996) 433–438.
[14] M. Gajhede, P. Osmark, F.M. Poulsen, H. Ipsen, J.N. Larsen, R.J.J. Van Neerven, C. Schou, H. Løwenstein, M.D. Spangfort, Nature Struct Biol 3 (1996) 1040–1045.
[15] W. Jost, H. Bak, K. Glund, P. Terpstra, J.J. Beintema, Eur J Biochem 198 (1991) 1–6.
[16] W. Zhao, J.J. Beintema, J. Hofsteenge, Eur J Biochem 219 (1994) 641–646.
[17] Allen G. Laboratory Techniques in Biochemistry and Molecular Biology. Sequencing of Proteins and Peptides. Amsterdam: Elsevier-North Holland 1981:142.
[18] J.D. Thompson, D.G. Higgins, T.J. Gibson, Nucleic Acids Res 22 (1994) 4673–4680.
[19] Swofford DL. PAUP (Phylogenetic Analysis Using Parsimony), Version 3.1.1. User's Manual. Washington, DC: Laboratory of Molecular Systematics, Smithsonian Institution, 1993.
[20] J. Felsenstein, Cladistics 5 (1989) 164–166.
[21] J. Hofsteenge, W.J. Weyer, P.A. Jekel, J.J. Beintema, Eur J Biochem 133 (1983) 91–108.
[22] S.A.J. Warner, R. Scott, J. Draper, Plant Mol Biol 19 (1992) 555–561.
[23] B.A. McClure, V. Haring, P.R. Ebert, M.A. Anderson, R.J. Simpson, F. Sakiyama, A.E. Clarke, Nature 342 (1989) 955–957.
[24] A.D. Richman, M.K. Uyenoyama, J.R. Kohn, Science 273 (1996) 1212–1216.
[25] J.J. Beintema, C. Schüller, M. Irie, A. Carsana, Prog Biophys Mol Biol 51 (1988) 165–192.
[26] H. Watanabe, H. Narumi, T. Inaba, K. Ohgi, M. Irie, J Biochem (Tokyo) 114 (1993) 800–807.
[27] M. Saraste, P.R. Sibbald, A. Wittinghofer, Trends Biochem Sci 15 (1990) 430–434.
[28] D.P. Matton, N. Brisson, Mol Plant-Microbe Interact 2 (1989) 325–331.