J. Biochem. **131**, 225–231 (2002)

Genomic Clones Encoding Two Isoforms of Pokeweed Antiviral Protein in Seeds (PAP-S1 and S2) and the *N*-Glycosidase Activities of Their Recombinant Proteins on Ribosomes and DNA in Comparison with Other Isoforms[1]

Eijiro Honjo, Danghong Dong, Hiroyuki Motoshima, and Keiichi Watanabe[2]

Department of Applied Biological Sciences, Saga University, Saga 840-8502

Received October 5, 2001; accepted November 27, 2001

Pokeweed antiviral proteins (PAPs) are single-chain ribosome-inactivating proteins (RIPs) isolated from several organs of *Phytolacca americana* (Pokeweed) that are characterized by their ability to depurinate not only ribosomes but also various nucleic acids. PAP-S is one of the isoforms found in seeds. In this study, we obtained three different genomic clones encoding two forms of PAP-S (here designated as PAP-S1 and PAP-S2) and α-PAP after PCR using a pair of degenerated primers based on the known N- and C-terminal amino acid sequences of PAP-S. The nucleotide sequences of the genomic clones contained no introns. The deduced amino acid sequences of PAP-S1 and PAP-S2, which showed 83% identity to each other, were found to correspond to sequences reported independently for PAP-S protein and cDNA, respectively, demonstrating that at least two forms of PAP-S actually exist in seeds of the same plant. The recombinant PAP-S1, PAP-S2, α-PAP, and PAP I (a form appearing in spring leaves) exhibit the same level of depurinating activity on rat ribosomes, while their efficiencies on *Escherichia coli* ribosomes and salmon sperm DNA differ substantially from one another in the order of PAP I > α-PAP > PAP-S1 > PAP-S2 and α-PAP > PAP I > PAP-S1 > PAP-S2. Structural comparisons suggest that the large difference in ribosome recognition between PAP-S1 (or S2) and PAP I is caused by the alteration of residues adjacent to the adenine-binding site.

Key words: *N*-glycosidase, *Phytolacca americana*, pokeweed antiviral protein, pokeweed seeds, ribosome-inactivating protein.

Pokeweed antiviral protein (PAP) from *Phytolacca americana* is a ribosome-inactivating protein (RIP) that enzymatically removes a single adenine base from a highly conserved "α-sarcin/ricin" loop of 28S (*1*) and 23S (*2*) rRNAs in eukaryotic and prokaryotic ribosomes, respectively. PAP belongs to type 1 RIPs, which are monomeric proteins with a molecular mass of around 30 kDa. In contrast, type 2 RIPs consist of an A chain that resembles type 1 RIP and a B chain that binds to galactose-containing receptors on the cell surface to facilitate penetration of the A chain into the cell (*3*). PAP, as well as other RIPs, have recently been reported to depurinate not only ribosomes and rRNA, but also DNA (*4*). PAP also exhibits a broad-spectrum of antiviral activity against plant and animal viruses, and has, therefore, attracted considerable interest for possible use in agricultural or medical fields (*3*).

Pokeweed produces several PAPs: PAP I (or PAP), PAP II and PAP-S, which appear in spring leaves, summer leaves and seeds, respectively (*3*). α-PAP is an unknown form found in a genomic DNA (*5*). The amino acid sequence of PAP I shows 76 and 74% identity to PAP-S and α-PAP respectively, while PAP II is only 33% identical to PAP I. The X-ray crystallographic structures of PAP I (*6*) and α-PAP (*7*) revealed that the overall structures of these proteins are homologous to the ricin toxin A-chain (RTA), which is a type 2 RIP, and that the residues that are perfectly conserved in all RIPs (Tyr72, Tyr123, Glu176, and Arg179 in PAP I) are located in the active site cleft. As the catalytic mechanism, it has been proposed that Tyr72 and Tyr123 sandwich the target adenine ring, and the side chain of Arg179 may protonate the N-3 atom of the adenine ring, whereas the negative charge of Glu176 may stabilize a positive oxocarbonium transition state (*6*). Other RIPs are also thought to function in a similar manner because of the resemblance of their active site configurations. In contrast, the substrate specificities of RIPs differ considerably from one another. For instance, PAP can act on both eukaryotic and prokaryotic ribosomes, but RTA is active only on eukaryotic ribosomes (*3*). Among PAP isoforms, PAP-S has been reported to be most effective in inhibiting protein syn-

---

[1] The sequence data of PAP-S1 and PAP-S2 have been deposited in the DDBJ/EMBL/GenBank databases with the accession numbers: AB071854 and AB071855, respectively. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan.

[2] To whom correspondence should be addressed. Tel/Fax: +81-952-28-8774, E-mail: watakei@cc.saga-u.ac.jp

Abbreviations: RIP, ribosome-inactivating protein; PAP, pokeweed antiviral protein; MBP, maltose-binding protein, RTA, ricin toxin A-chain.

thesis in rabbit reticulocyte lysate (*8, 9*).

We have shown that PAP-S purified from seeds efficiently inactivates tobacco ribosomes and inhibits virus multiplication in tobacco cells (*10*), suggesting a possible use of the PAP-S gene to produce transgenic plants against pathogens. PAP-S also effectively inhibits the replication of human immunodeficiency virus 1 in infected peripheral blood mononuclear cells (*11*). We think that elucidating structural factors affecting the substrate specificity of PAP will facilitate the protein engineering of PAP to modify its function to one more suitable for agricultural or therapeutic applications. Since the amino acid sequence of PAP-S is very similar to that of PAP I, these proteins seem to be suitable for comparative studies on the structure–function relationships of PAP. However, quantitative data comparing the *N*-glycosidase activities for various substrates among PAP isoforms have not been available. Furthermore, the amino acid sequence of PAP-S deduced from the cDNA (*9*) is not identical to that determined by direct protein sequencing (*12*) (83% identity), and this difference has been speculated to be due to the presence of different forms of PAP-S in seeds, to heterogeneity of PAP-S in pokeweed variants, or to sequencing errors (*9*), but the facts remain unclear.

The main purpose of this study is to explore the structural factors affecting the substrate specificity of PAP by comparing the structures and activities with various substrates among PAP isoforms. Since we have already established the expression and purification system for the recombinant PAP I (*13*), we tried to clone the genomic DNA fragment encoding the mature PAP-S. Here, we report our findings that two forms of PAP-S (designated as PAP-S1 and PAP-S2) actually exist in the seeds of the same plant, and that the recombinant PAP-S1, PAP-S2, PAP I, and α-PAP differ substantially in substrate specificity. The structural comparison among PAP isoforms allowed us to demonstrate the amino acid residues that may be involved in the recognition of different substrates.

MATERIALS AND METHODS

Materials—Ex Taq DNA polymerase was purchased from Takara Shuzo (Kyoto). Restriction endonucleases, T4 DNA ligase, β-agarase I and agarose L were from Nippon Gene (Tokyo). Salmon sperm DNA from Wako Pure Chemicals (Tokyo) was purified by phenol extraction, ethanol precipitation and treatment with DNase-free RNase A (Boehringer) followed by a second phenol extraction and ethanol precipitation. A Thermo Sequence II dye terminator cycle sequence kit and Mono-S column were from Amersham Pharmacia Biotech. The pBluescript vector used for cloning and *Escherichia coli* strain XL1-Blue were from Stratagene. Plasmid pMAL-p2, amylose resin, and factor Xa were from New England BioLabs. The oligonucleotides used for cloning and site-directed mutagenesis were synthesized by Spec Oligo Service (Tokyo). PAP I was expressed in *E. coli* and purified as previously reported (*13*). Rat liver polysomes were prepared by the established method (*14*). *E. coli* ribosomes were prepared from strain XL1-Blue (*15*).

Amplification of DNA by PCR—Genomic DNA was prepared from spring leaves of *Phytolacca americana* by phenol extraction. PAP-specific DNA fragments were amplified from the genomic DNA by PCR using Ex Taq polymerase, *ps-n* primer (5′-GGGGGATCCAT(ACT)AA(CT)AC(ACGT)AT(ACT)AC(ACGT)TT(CT)GA(CT)GC-3′) designed according to the N-terminal sequence of mature PAP-S with a *Bam*HI site, and *ps-c* primer (5′-CCCAAGCTTCA(ACGT)GT(ACGT)GC(CT)TG(AG)CA(ACGT)GT(ACGT)CC-3′) designed according to the C-terminal sequence with a stop codon and a *Hin*dIII site. The reaction mixture contained 500 ng genomic DNA, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2 mM $MgCl_2$, 100 pmol of primers, 200 μM of each dNTP, and 1.25 units of Ex Taq polymerase in a total volume of 100 μl. The reaction was performed under the following incubation conditions: 3 min at 94°C; 30 cycles of 1 min at 94°C, 2 min at 50°C for annealing, 3 min at 72°C for extension; 10 min at 72°C for a final extension. The amplified DNA fragments were separated by agarose gel electrophoresis. A fragment with the expected size (about 800 bp) was extracted and used as a template for the second PCR, the conditions of which were the same as for the first PCR with the exception that the annealing temperature was 48°C. This PCR mainly amplified a fragment of about 800 bp, which was purified by agarose gel electrophoresis, digested with *Bam*HI and *Hin*dIII, and cloned into pBluescript vector using *E. coli* strain XL1-Blue as a host cell. These DNA manipulations followed standard methods (*16*).

DNA Sequencing—The nucleotide sequences of inserts cloned into plasmid DNA were determined on both strands using a Thermo Sequence II dye terminator cycle sequencing kit and a 373A DNA sequencer (Applied Biosystems). The double stranded pBluescript containing the DNA insert was used as a template with primers for M13, M13 reverse, or oligonucleotides synthesized according to the sequences determined.

Mutagenesis—Site-directed mutagenesis was performed by PCR (*17*) using a mutagenic primer, two flanking primers, and pBluescript containing a DNA fragment encoding mature PAP-S1 as a template. The N-terminal primer with a *Bam*HI site, 5′-GGGATCCATCAATACGATAACGTTCGACG-3′, and C-terminal primer with a stop codon and *Hin*dIII site, 5′-GGAAGCTTCAAGTGGCTTGGCAGGTACC-3′, were designed according to the determined DNA sequence of mature PAP-S1. The primer *ps-bam* (5′-GTCGTAGGGGTCCGAATAGC-3′) was used for alteration of the *Bam*HI site in the PAP-S1 DNA without any change in the encoded amino acid sequence. The PCR fragment containing the point mutation was cloned into pBluescript and sequenced to ensure that no unexpected mutation had occurred.

Expression and Purification of Recombinant Mature PAP—The expression of each PAP isoform was performed as described previously (*13*). The DNA fragments encoding the mature regions of PAP isoforms were inserted between the *Bam*HI and *Hin*dIII sites of the expression vector pMAL-p2 and introduced into *E. coli* strain XL1-Blue. The transformant was grown at 37°C in 1 liter of LB medium containing ampicillin (100 μg/ml) and tetracycline (12 μg/ml) to an OD of 1.0 at 600 nm, and then incubated with 0.35 mM isopropyl-β-D-thiogalactopyranoside for 5 h at 25°C. The PAP fusion with maltose-binding protein was extracted from the periplasmic fractions of harvested bacterial cells, and then purified by affinity chromatography on an amylose resin column. After digestion of the fusion with factor Xa, PAP was purified by Mono-S column chromatography.

Determination of N-Glycosidase Activity—The RNA *N*-glycosidase activity of PAP isoforms on rat liver and *E. coli* ribosomes was assayed by measuring adenine released from ribosomes by the action of PAP as described previously (*18*). The released adenine was recovered as an ethanol-soluble fraction after precipitation of the ribosomes with 50% ethanol, converted to a fluorescing derivative, 1,N^6-ethenoadenine, by reaction with chloroacetaldehyde, and then analyzed by reverse-phase HPLC monitoring the fluorescence of the eluate at excitation and emission wavelengths of 280 and 400 nm, respectively. The protein concentration was measured using bovine serum albumin as a standard according to Ref. *19*.

The DNA *N*-glycosidase activity of PAP isoforms was assayed by measuring adenine released from salmon sperm DNA. The DNA (20 μg) was incubated with 1.6 μg of each PAPs in 40 μl of 5 mM Tris-HCl, pH 7.5, 50 mM KCl, 5 mM Mg acetate, and 1 mM dithiothreitol for 2 h at 37°C. After ethanol precipitation, the released adenine was recovered as an ethanol-soluble fraction. The fraction was dried under vacuum and then dissolved in 100 μl of 10 mM potassium phosphate buffer, pH 6.0. The sample (20 μl) was subjected to reverse-phase HPLC on a column of Wakosil 5C18 (4.6 × 250 mm, Wako Pure Chemicals) and eluted with 18% methanol in 10 mM potassium phosphate buffer, pH 6.0. Adenine was detected by absorbance at 260 nm.

Construction of the Model—The model of PAP-S1 was constructed with the program MODELLER 4 (*20*) implemented on a Silicon Graphics workstation using the X-ray structure of PAP I (*6*) as a template. Alignment of the amino acid sequences was clearly made because of high identity (76%) of the sequences and only one deletion in a loop of PAP-S1. Two intramolecular disulfide bonds were initially introduced into the model as chemically identified (*12*). The model was evaluated by PROCHECK (*21*) and Verify3D (*22*).

RESULTS AND DISCUSSION

Amplification and Cloning of DNAs Encoding PAP Isoforms—The DNA fragment of about 800 bp expected to encode the mature PAP-S was amplified from the genomic DNA by PCR using a pair of degenerated primers (*ps-n* and *ps-c*) based on the N-terminal and C-terminal sequences of PAP-S (*12*). This fragment was designated as PF1 (Fig. 1). Digestion with *Hin*dIII did not cause fragmentation of PF1, but the following digestion with *Bam*HI produced an additional two fragments of about 600 bp (PF2) and 200 bp (PF3) (Fig. 1). This suggested that the original PF1 fragment was a mixture of at least two species of the PAP gene, one of which was considered to be a new gene, because the DNA sequences so far reported for PAP I, α-PAP, and PAP-S do not have a *Bam*HI site (*5, 9, 23*). All of the resulting DNA fragments were, therefore, ligated into pBluescript vector between the *Bam*HI and *Hin*dIII sites or at the *Bam*HI site alone. The PF1 and PF2 fragments were cloned between the *Bam*HI and *Hin*dIII sites of the vector and the PF3 fragment was at the *Bam*HI site. DNA sequencing revealed that the cloned PF1 fragment corresponded to either the cDNA for PAP-S (*9*) or the genomic clone of α-PAP (*5*), although four alterations (Ala11 to Ser, Met126 to Leu, Asn136 to Ser, and Thr261 to Ala) and one alteration (Ile209 to Thr) in the deduced amino acid sequences of the present PAP-S and α-PAP were seen, respectively. On the other hand, the cloned PF3 and PF2 fragments were found to encode the N-terminal side (78 amino acids) and the residual C-terminal side (183 amino acids) of mature PAP-S, respectively (Fig. 2A), and the deduced amino acid sequence was identical to that determined by direct protein sequencing (*12*), except for a single substitution of Glu136 to Gln. The genomic clones encoding two different isoforms of mature PAP-S did not contain any intron as PAP I (*23*) and α-PAP (*5*) (Fig. 2, A and B).

The amino acid sequence of PAP-S deduced from fragments PF3 and PF2 was 83% identical to that deduced from fragment PF1. Here we designate the former as PAP-S1 and the latter as PAP-S2. The existence of two different isoforms of PAP in the seeds of the same plant has not so far been clear, because the difference in the sequences reported independently for the PAP-S protein (*12*) and cDNA (*9*) has been speculated to be due to the heterogeneity of PAP-S in pokeweed variants, to sequencing error, or to the presence of different forms of PAP-S in seeds (*9*). Our results confirm that at least two different forms of PAP-S (PAP-S1 and PAP-S2) actually exist in seeds from the same plant.

The *Bam*HI site inside the PAP-S1 DNA sequence, was not favorable to the construction of the expression vector using pMAL-p2. It was, therefore, modified by site-directed mutagenesis. The pBluescript vector containing the overall DNA sequence of mature PAP-S1 was first constructed by inserting the PF3 fragment into the *Bam*HI site of pBluescript/PF2 and then used as a template for PCR for mutagenesis. By DNA sequencing, it was confirmed that a *Bam*HI site within the PAP-S1 DNA was disrupted by substitution of a single base (GGATCC to GGACCC) without any alterations in the encoded amino acid sequence.

Expression and Purification of PAP Isoforms—Purification of the recombinant PAP-S1, PAP-S2, and α-PAP following the expression in *E. coli* as a fusion with maltose-binding protein resulted in a final yield of 0.02 to 0.05 mg/liter of culture. These yields were much lower than that of PAP I (4.5 mg/liter of culture). This is accounted for by the fact that the expression of PAP-S1, PAP-S2, and α-PAP caused more severe growth-inhibition of the host cells than the expression of PAP I. Each of the four PAP isoforms showed a single band on SDS-PAGE (12.5%), with an identical mobility corresponding to a molecular mass of 30 kDa (data not shown).

N-Glycosidase Activity on Ribosomes and on DNA—The

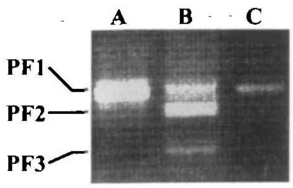


Fig. 1. **Purified PCR products digested with *Hin*dIII and *Bam*HI.** The second PCR product (500 ng) purified by agarose gel electrophoresis was digested with 20 units of *Hin*dIII and then *Bam*HI for 2 h at 37°C. The digests were separated in a 1.5% agarose gel, stained with ethidium bromide, and visualized under an ultraviolet lamp. Lane A, digest with *Hin*dIII; B, digest with *Hin*dIII and *Bam*HI; C, control DNA fragment encoding mature PAP I (about 800 bp).

A

```
1      ATCAATACGATAACGTTCGACGCTGGAAATGCAACCATTAACAAG
1       I  N  T  I  T  F  D  A  G  N  A  T  I  N  K

46     TATGCTACCTTTATGGAATCTCTTCGTAATGAAGCGAAAGATCCA
16      Y  A  T  F  M  E  S  L  R  N  E  A  K  D  P

91     AGTTTAAAGTGTTATGGAATACCAATGTTGCCCAATACTAATTCA
31      S  L  K  C  Y  G  I  P  M  L  P  N  T  N  S

136    ACCATCAAGTACTTGTTGGTTAAGCTCCAAGGTGCAAGCCTAAAA
46      T  I  K  Y  L  L  V  K  L  Q  G  A  S  L  K

181    ACCATCACACTAATGCTAAGACGAAACAACTTATATGTGATGGGC
61      T  I  T  L  M  L  R  R  N  N  L  Y  V  M  G

226    TATTCGGATCCCTACGACAATAAGTGTCGTTACCATATCTTTAAT
76      Y  S  D  P  Y  D  N  K  C  R  Y  H  I  F  N

271    GATATTAAAGGCACTGAATACAGTGATGTGGAGAATACTCTTTGC
91      D  I  K  G  T  E  Y  S  D  V  E  N  T  L  C

316    CCAAGTTCAAATCCTCGTGTTGCAAAACCCATTAACTACAATGGC
106     P  S  S  N  P  R  V  A  K  P  I  N  Y  N  G

361    CTATATCCAACTTTGGAAAAAAAAGCAGGAGTAACCTCAAGAAAT
121     L  Y  P  T  L  E  K  K  A  G  V  T  S  R  N

406    CAAGTCCAACTAGGAATTCAAATACTCAGCAGTGACATTGGAAAA
136     Q  V  Q  L  G  I  Q  I  L  S  S  D  I  G  K

451    ATCTCTGGACAGGGCTCGTTCACTGAAAAAATCGAGGCCAAATTC
151     I  S  G  Q  G  S  F  T  E  K  I  E  A  K  F

496    CTGCTTGTAGCCATTCAAATGGTGTCAGAAGCAGCGCGATTCAAG
166     L  L  V  A  I  Q  M  V  S  E  A  A  R  F  K

541    TACATAGAGAATCAGGTGAAGACTAATTTTAACAGAGATTTCTCC
181     Y  I  E  N  Q  V  K  T  N  F  N  R  D  F  S

586    CCTAATGACAAAGTACTTGACTTGGAGGAGAACTGGGGTAAGATC
196     P  N  D  K  V  L  D  L  E  E  N  W  G  K  I

631    TCTACGGCAATTCACAATTCCAAGAATGGAGCTTTACCAAAACCT
211     S  T  A  I  H  N  S  K  N  G  A  L  P  K  P

676    CTAGAGCTAAAAAATGCAGACGGTACTAAGTGGATAGTGCTTAGA
226     L  E  L  K  N  A  D  G  T  K  W  I  V  L  R

721    GTGGATGAAATCAAACCTGATGTGGGACTCCTTAACTATGTTAAT
241     V  D  E  I  K  P  D  V  G  L  L  N  Y  V  N

766    GGTACCTGCCAAGCCACT
256     G  T  C  Q  A  T
```

B

```
1      ATTAACACCTTTACCTTTGATGCGGGCAACAGCACCATTAACAAA
1       I  N  T  F  T  F  D  A  G  N  S  T  I  N  K

46     TATGCGACCTTTATGGAAAGCCTGCGTAACCAGGCGAAAGATCCG
16      Y  A  T  F  M  E  S  L  R  N  Q  A  K  D  P

91     AAACTGAAATGCTATGGCATTCCGATGCTGCCGGATACCAACAGC
31      K  L  K  C  Y  G  I  P  M  L  P  D  T  N  S

136    ACCCCGAAATATCTGCTGGTGAAACTGCAGGGCGCGAACCTGAAA
46      T  P  K  Y  L  L  V  K  L  Q  G  A  N  L  K

181    ACCATTACCCTGATGCTGCGTCGTAACAACCTGTATGTGATGGGC
61      T  I  T  L  M  L  R  R  N  N  L  Y  V  M  G

226    TATAGCGATCCGTTTAACGGCAACAAATGCCGTTATCATATTTTT
76      Y  S  D  P  F  N  G  N  K  C  R  Y  H  I  F

271    AACGATATTACCAGCACCGAACGTACCGATGTGGAAAACACCCTG
91      N  D  I  T  S  T  E  R  T  D  V  E  N  T  L

316    TGCAGCAGCAGCAGCAGCCGTGTGGCGATGAGCATTAACTATAAC
106     C  S  S  S  S  S  R  V  A  M  S  I  N  Y  N

361    AGCCTGTATCCGACCCTGGAAAAAAAAGCGGAAGTGAACAGCCGT
121     S  L  Y  P  T  L  E  K  K  A  E  V  N  S  R

406    AGCCAGGTGCAGCTGGGCATTCAGATTCTGAGCAGCGATATTGGC
136     S  Q  V  Q  L  G  I  Q  I  L  S  S  D  I  G

451    AAAATTAGCGGCGTGGATAGCTTTCCGGTGAAAACCGAAGCGTTT
151     K  I  S  G  V  D  S  F  P  V  K  T  E  A  F

496    TTTCTGCTGGTGGCGATTCAGATGGTGAGCGAAGCGGCGCGTTTT
166     F  L  L  V  A  I  Q  M  V  S  E  A  A  R  F

541    AAATATATTGAAAACCAGGTGAAAACCAACTTTAACCGTGCGTTT
181     K  Y  I  E  N  Q  V  K  T  N  F  N  R  A  F

586    TATCCGGATCCGAAAGTGATTAACCTGGAAGAAAAATGGGGCAAA
196     Y  P  D  P  K  V  I  N  L  E  E  K  W  G  K

631    ATTAGCGAAGCGATTCATAACGCGAAAAACGGCGCGCTGCCGAAA
211     I  S  E  A  I  H  N  A  K  N  G  A  L  P  K

676    CCGCTGGAACTGGTGGATGCGAAAGGCACCAAATGGATTGTGCTG
226     P  L  E  L  V  D  A  K  G  T  K  W  I  V  L

721    CGTGTGGATGAAATTAACCGTGATGTGGCGCTGCTGAAATATGTG
241     R  V  D  E  I  N  R  D  V  A  L  L  K  Y  V

766    AACGGCACCTGCCAGGCGACC
256     N  G  T  C  Q  A  T
```

Fig. 2. **Nucleotide sequences of genomic clones and deduced amino acid sequences of mature PAP-S1 (A) and PAP-S2 (B).** The underlined residues in the DNA sequence of PAP-S1 indicates the *Bam*HI site.

RNA *N*-glycosidase activities on rat liver and *E. coli* ribosomes were tested for recombinant PAP-S1, PAP-S2, PAP I, and α-PAP. Figure 3 illustrates adenine release from ribosomes after reaction with various concentrations of PAP isoforms. Since all PAP isoforms had much lower activity on *E. coli* ribosomes than on rat liver ribosomes, the reaction with *E. coli* ribosomes was performed at 37°C for 60 min instead of 10 min with rat liver ribosomes. All PAP isoforms were able to release one mole of adenine from one mole of both rat liver and *E. coli* ribosomes, that is, from a highly conserved α-sarcin/ricin loop, a common target of the enzymatic action of RIPs (*1, 2*). The concentrations of PAP isoforms required for 50% of the maximum release of adenine from ribosomes (RC_{50}) are listed in Table I. PAP-S1 and PAP-S2 had the same activity on rat liver ribosomes, which was also virtually the same level as for PAP I and α-PAP. On the other hand, the activities of PAP isoforms on *E. coli* ribosomes were found to be substantially different from one another in the order of PAP I > α-PAP > PAP-S1 > PAP-S2. PAP-S1 was 3.6-fold more effective than PAP-S2, but 11-fold and 60-fold less active than α-PAP and PAP I, respectively. This is the first quantitative evidence of large differences in activity on prokaryotic ribosomes among PAP isoforms. Our results showing that the activity of PAP-S2 on mammal ribosomes is similar to that of PAP I is not consistent with a previous report that PAP-S (corresponding to PAP-S2 in this study) is more efficient than PAP I in inhibiting cell-free protein synthesis using a rabbit reticulocyte system (*9*). This discrepancy could be due to the difference in the assay systems used, or to four amino acid substitutions observed in PAP-S2. Our assay is just for *N*-glycosidase activity on ribosomes, while inhibition of cell-free protein synthesis might reflect the total effect on ribosomes, tRNA and mRNA, because PAP can act on a wide variety of nucleic acids as an *N*-glycosidase (*4*).

Adenine release from salmon sperm DNA by the action of PAPs was also tested at physiological pH (Table I). The efficiency on DNA was in the order of α-PAP > PAP I > PAP-S1 > PAP-S2. The amount of adenine released by PAP-S2 was 9.3-fold, 32-fold and 82-fold lower than those of PAP-S1, PAP I, and α-PAP, respectively. The order of the DNA *N*-glycosidase activity correlates fairly well with the RNA *N*-glycosidase activity on *E. coli* ribosomes except in the case of α-PAP and PAP I. The present results showing a lower efficiency of PAP-S1 on DNA than PAP I is in agreement with a previous report (*4*), in which herring sperm DNA was reacted at pH 4.0 with PAP isoforms purified from plant sources, although the difference observed here at pH 7.5 is clearer than the previous result.

The observed significant distinction in substrate specificity among PAP isoforms may reflect their different physiological substrates or different roles in plants such as a defensive role against pathogens or a metabolic role (*3, 24-26*). For the antiviral action, we have previously shown that PAP (PAP-S1 in this study) selectively enters virus-infected tobacco cells and prevents viral multiplication by inactivating host ribosomes (*10*). However, the direct action of PAP on viral RNA and/or virus-derived DNA can not be excluded. Moreover, a recent study on transgenic tobacco expressing a non-toxic PAP mutant has demonstrated that PAP-mediated resistance against both viral and fungal infection is caused by the activation of a novel salicylic acid-independent, stress-associated signal transduction pathway without depurination of rRNA (*25*). Since the actual target of PAP for pathogen resistance has not yet been established, testing each PAP isoform and its mutants with different specificities may be valuable for developing transgenic plants more resistant to pathogens.

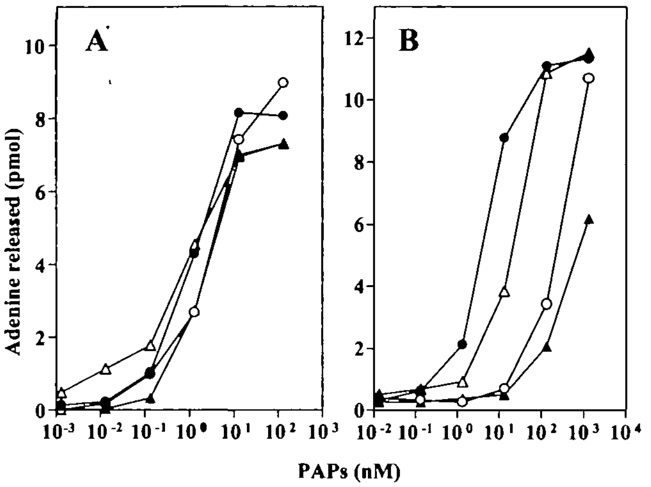
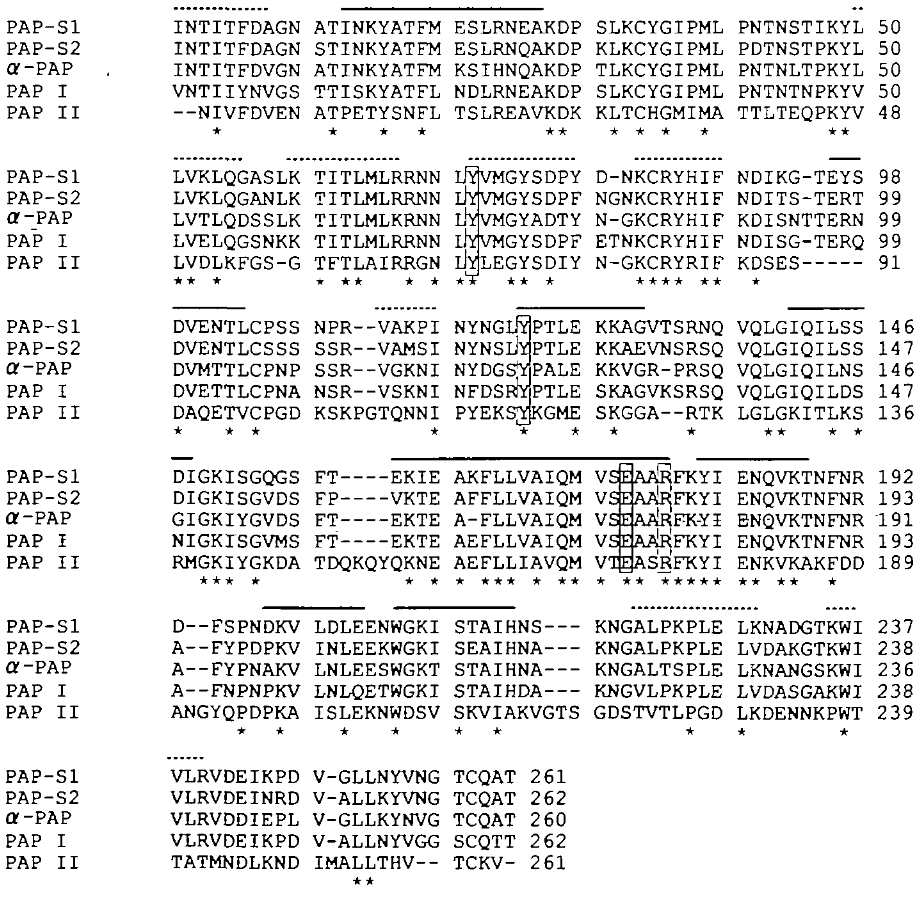


Fig. 3. **RNA *N*-glycosidase activities of PAP isoforms on rat liver (A) and *E. coli* (B) ribosomes.** Rat liver ribosomes (10 pmol) and *E. coli* ribosomes (9.4 pmol), in 5 mM Tris-HCl, pH 7.5, 50 mM KCl, 5 mM Mg acetate, and 1 mM dithiothreitol, were reacted with increasing amounts of PAP I (●), PAP-S1 (○), PAP-S2 (▲), and α-PAP (△) at 37°C for 10 min and for 1 h, respectively. Adenine released from ribosomes was measured after conversion to a fluorescing derivative, 1,N^6-ethenoadenine.

Structural Comparison—In spite of the substantial difference in the activity on prokaryotic ribosomes and salmon sperm DNA among PAP isoforms, the amino acid sequences of PAP-S1, PAP-S2, PAP I, and α-PAP are very similar to one another, showing 73 to 83% identity (Fig. 4). Only PAP II has a sequence that is much different from

TABLE I. **Depurinating activity of PAP isoforms.**

PAP isoform	Concentrations of PAP required for 50% of the maximum release of adenine (RC_{50})		Adenine released from DNA (pmol)
	Rat liver ribosomes (nM)	*E. coli* ribosomes (nM)	
PAP-S1	3.2	280	56
PAP-S2	3.6	1,000	6
PAP I	1.5	4.7	190
α-PAP	1.3	25	490



```
PAP-S1          INTITFDAGN ATINKYATFM ESLRNEAKDP SLKCYGIPML PNTNSTIKYL 50
PAP-S2          INTITFDAGN STINKYATFM ESLRNQAKDP KLKCYGIPML PDTNSTPKYL 50
α-PAP           INTITFDVGN ATINKYATFM KSIHNQAKDP TLKCYGIPML PNTNLTPKYL 50
PAP I           VNTIIYNVGS TTISKYATFL NDLRNEAKDP SLKCYGIPML PNTNTNPKYV 50
PAP II          --NIVFDVEN ATPETYSNFL TSLREAVKDK KLTCHGMIMA TTLTEQPKYV 48
                   *        *   *  *         **   * * *  *         **

PAP-S1          LVKLQGASLK TITLMLRRNN LYVMGYSDPY D-NKCRYHIF NDIKG-TEYS 98
PAP-S2          LVKLQGANLK TITLMLRRNN LYVMGYSDPF NGNKCRYHIF NDITS-TERT 99
α-PAP           LVTLQDSSLK TITLMLKRNN LYVMGYADTY N-GKCRYHIF KDISNTTERN 99
PAP I           LVELQGSNKK TITLMLRRNN LYVMGYSDPF ETNKCRYHIF NDISG-TERQ 99
PAP II          LVDLKFGS-G TFTLAIRRGN LYLEGYSDIY N-GKCRYRIF KDSES----- 91
                ** *       * **   * * **  ** *      **** **  *

PAP-S1          DVENTLCPSS NPR--VAKPI NYNGLYPTLE KKAGVTSRNQ VQLGIQILSS 146
PAP-S2          DVENTLCSSS SSR--VAMSI NYNSLYPTLE KKAEVNSRSQ VQLGIQILSS 147
α-PAP           DVMTTLCPNP SSR--VGKNI NYDGSYPALE KKVGR-PRSQ VQLGIQILNS 146
PAP I           DVETTLCPNA NSR--VSKNI NFDSRYPTLE SKAGVKSRSQ VQLGIQILDS 147
PAP II          DAQETVCPGD KSKPGTQNNI PYEKSYKGME SKGGA--RTK LGLGKITLKS 136
                *   * *             *      *   *  *     *     **   * *

PAP-S1          DIGKISGQGS FT----EKIE AKFLLVAIQM VSEAARFKYI ENQVKTNFNR 192
PAP-S2          DIGKISGVDS FP----VKTE AFFLLVAIQM VSEAARFKYI ENQVKTNFNR 193
α-PAP           GIGKIYGVDS FT----EKTE A-FLLVAIQM VSEAARFKYI ENQVKTNFNR 191
PAP I           NIGKISGVMS FT----EKTE AEFLLVAIQM VSEAARFKYI ENQVKTNFNR 193
PAP II          RMGKIYGKDA TDQKQYQKNE AEFLLIAVQM VTEASRFKYI ENKVKAKFDD 189
                  *** *           * * * *** * ** * ** ***** ** **  *

PAP-S1          D--FSPNDKV LDLEENWGKI STAIHNS--- KNGALPKPLE LKNADGTKWI 237
PAP-S2          A--FYPDPKV INLEEKWGKI SEAIHNA--- KNGALPKPLE LVDAKGTKWI 238
α-PAP           A--FYPNAKV LNLEESWGKT STAIHNA--- KNGALTSPLE LKNANGSKWI 236
PAP I           A--FNPNPKV LNLQETWGKI STAIHDA--- KNGVLPKPLE LVDASGAKWI 238
PAP II          ANGYQPDPKA ISLEKNWDSV SKVIAKVGTS GDSTVTLPGD LKDENNKPWT 239
                     *  *    *   *    *  *              *   *       *

PAP-S1          VLRVDEIKPD V-GLLNYVNG TCQAT 261
PAP-S2          VLRVDEINRD V-ALLKYVNG TCQAT 262
α-PAP           VLRVDDIEPL V-GLLKYNVG TCQAT 260
PAP I           VLRVDEIKPD V-ALLNYVGG SCQTT 262
PAP II          TATMNDLKND IMALLTHV-- TCKV- 261
                              **
```

Fig. 4. **Alignment of deduced amino acid sequences of mature PAP isoforms.** The sequences of five PAP isoforms were aligned and gaps (—) were introduced to maximize homology using the ClustalX program (*27*). Identical amino acids are marked by asterisks (*). Amino acid residues possibly involved in catalytic activity are boxed. Regions of α helix (——) and β strands (——) are indicated above the sequences.

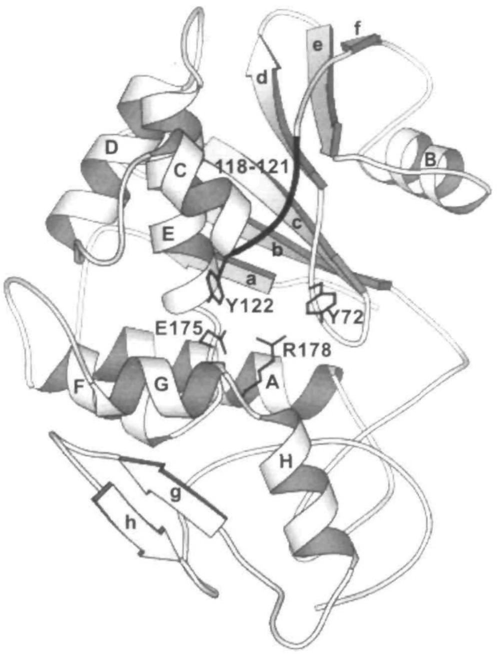


Fig. 5. **The model of PAP-S1.** The model of PAP-S1 was constructed based on the X-ray structure of PAP I (*6*). The α helices are labeled A to H and the β strands are a to h. The side chains of putative active residues are displayed. Residues 118 to 121, indicated by a black string, are all different from those in PAP I, which is likely to cause the difference in ribosome recognition between these isoforms. The structure was drawn with the program Molscript (*28*).

those of other PAP isoforms. The X-ray crystal structure of PAP I has been solved (*6*) and some residues in the putative active site cleft that are conserved in all RIPs are proposed to be involved in the *N*-glycosidase activity. One can easily imagine that PAP-S1 and PAP-S2 have three dimensional structures very similar to that of PAP I because of the high identity of the overall sequences and the four conserved cysteine residues forming two intramolecular disulfide bonds: one between residues 34 and 258 and the other between residues 84 and 105 as chemically identified in PAP-S1 (*12*). An energetically favorable model of PAP-S1 was constructed based on the structure of PAP I (Fig. 5). For the catalytic mechanism of PAP-S1, it can be considered by analogy with PAP I that the adenine ring of the substrate is sandwiched between Tyr72 and Tyr122 through a rotation about the Cα–Cβ bond of Tyr72, with hydrolysis of the *N*-glycosidic bond by the concerted action of Glu175 and Arg178.

A question of a greater interest concerning protein engineering of PAP is which amino acid alterations cause the observed significant difference in substrate specificity between PAP-S1 (or S2) and PAP I. The enzymatic activities of PAP-S1 and PAP-S2 on *E. coli* ribosomes are the same, but 60-fold and 210-fold lower than that of PAP I, respectively (Fig. 3B and Table I). Thus the residues that are conserved between PAP-S1 and PAP-S2 but different in PAP I may bring about this difference in recognition of prokaryotic ribosomes. Alterations of four successive amino acid residues between PAP-S1 and PAP I are seen at residues 5 to 8 and 118 to 121 (numbering of PAP-S1) (Fig. 4). These residues in PAP-S2 are identical to those in PAP-S1 except for a single substitution of Ser for Gly120. Residues 5 to 8 are located just behind the active site cleft, but residues 118 to 121 are adjacent to the adenine-binding site between Tyr72 and Tyr122 (Fig. 5). Therefore, the latter residues (118-YNGL in PAP-S1, 119-YNSL in PAP-S2, 119-FDSR in PAP I) seem to be prime candidates for structural factors affecting the recognition of prokaryotic ribosomes. Indeed, the side chain of Arg122 in PAP I (corresponding to Leu121 in PAP-S1) has been shown to interact with the phosphate group of a substrate analog, formycin 5′-monophosphate by X-ray crystallography (*6*). This interaction be absent from both PAP-S1 and PAP-S2 because of the substitution of Leu for Arg, possibly causing a decrease in activity on prokaryotic ribosomes. This assumption is consistent with the lower activity of α-PAP on prokaryotic ribosomes than that of PAP I, because α-PAP has a Ser residue corresponding to Arg122 of PAP I. The recent alanine scanning mutagenesis of PAP I has suggested that polar side chains located at the active site cleft (48-KY, 67-RRNN, 90-FND) also participate in ribosome recognition (*29*). However, these residues are perfectly conserved in the three isoforms (PAP-S1, PAP-S2, and PAP-I), indicating no contribution to the observed difference in the recognition of prokaryotic ribosomes. Our chemical modification studies on RTA and luffin-a (a RIP from the seeds of *Luffa cylindrica*) have shown that residues far from the catalytic site are also involved in activity (*18, 30, 31*). The residues identified as participating in activity are Arg235 in RTA and Lys231 in luffin-a, which correspond to Ala231 and Arg240 in PAP-S1. Ala231 is situated in the loop between β strands g and h, and Arg240 is at the c-terminus of β strand h. The residues around Arg240 are identical among four isoforms (PAP-S1, PAP-S2, PAP-I, and α-PAP), while those around Ala231 show relatively large differences among these isoforms. Thus the residues in the loop connecting β strands g and h are also worth examining. Concerning the activity on DNA, the manner of a recognition by PAP may differ somewhat from that of ribosome recognition, because α-PAP instead of PAP I showed the highest efficiency in depurinating salmon sperm DNA among the isoforms tested (Table I). α-PAP has a unique positive side chain at Lys90 in residues 89-FKD located at the active site cleft, which correspond to the residues FND conserved in other three isoforms, and have been proposed to be involved in ribosome recognition (*29*) as described above. The possibility of the contribution of this characteristic positive side chain in α-PAP to DNA recognition is also worth examining. We are currently analyzing site-directed mutants of these isoforms to investigate the relationship between structure and substrate specificity, which would assist protein engineering aiming at making a PAP more suitable for agricultural or therapeutic applications.

REFERENCES

1. Endo, Y., Mitsui, K., Motizuki, M., and Tsurugi, K. (1987) The mechanism of action of ricin and related toxic lectins on eukaryotic ribosomes. *J. Biol. Chem.* **262**, 5908–5912
2. Hartley, M.R., Legname, G., Osborn, R., Chen, Z., and Lord, J.M. (1991) Single-chain ribosome inactivating proteins from plants depurinate *Escherichia coli* 23S ribosomal RNA. *FEBS Lett.* **290**, 65–68
3. Barbieri, L., Battelli, M.G., and Stirpe, F. (1993) Ribosome-inactivating proteins from plants. *Biochim. Biophys. Acta* **1154**,

237–282

4. Barbieri, L., Valbonesi, P., Bonora, E., Gorini, P., Bolognesi, A., and Stirpe, F. (1997) Polynucleotide:adenosine glycosidase activity of ribosome-inactivating proteins: effect on DNA, RNA and poly(A). *Nucleic Acids Res.* **25**, 518–522
5. Kataoka, J., Habuka, N., Masuta, C., Miyano, M., and Koiwai, A. (1992) Isolation and analysis of a genomic clone encoding a pokeweed antiviral protein. *Plant Mol. Biol.* **20**, 879–886
6. Monzingo, A.F., Collins, E.J., Ernst, S.R., Irvin, J.D., and Robertus, J.D. (1993) The 2.5 Å structure of pokeweed antiviral protein. *J. Mol. Biol.* **233**, 705–715
7. Ago, H., Kataoka, J., Tsuge, H., Habuka, N., Inagaki, E., Noma, M., and Miyano, M. (1994) X-ray structure of a pokeweed antiviral protein, coded by a new genomic clone, at 0.23 nm resolution. *Eur. J. Biochem.* **225**, 369–374
8. Barbieri, L., Aron, G.M., Irvin, J.D., and Stirpe, F. (1982) Purification and partial characterization of another form of the antiviral protein from the seeds of *Phytolacca americana* L. (pokeweed). *Biochem. J.* **203**, 55–59
9. Poyet, J-L. and Hoeveler, A. (1997) cDNA cloning and expression of pokeweed antiviral protein from seeds in *Escherichia coli* and its inhibition of protein synthesis in vitro. *FEBS Lett.* **406**, 97–100
10. Watanabe, K., Kawasaki, T., Sako, N., and Funatsu, G. (1997) Actions of pokeweed antiviral protein on virus-infected protoplasts. *Biosci. Biotechnol. Biochem.* **61**, 994–997
11. Olson, M.C., Ramakrishnan, S., and Anand, R. (1991) Ribosomal inhibitory proteins from plants inhibit HIV-1 replication in acutely infected peripheral blood mononuclear cells. *AIDS Res. Hum. Retroviruses* **7**, 1025–1030
12. Kung, S-S., Kimura, M., and Funatsu, G. (1990) The complete amino acid sequence of antiviral protein from the seeds of pokeweed (*Phytolacca americana*). *Agric. Biol. Chem.* **54**, 3301–3318
13. Honjo, E. and Watanabe, K. (1999) Expression of mature pokeweed antiviral protein with or without C-terminal extrapeptide in *Escherichia coli* as a fusion with maltose-binding protein. *Biosci. Biotechnol. Biochem.* **63**, 1291–1294
14. Staehelin, T. and Falvey, A.K. (1971) Isolation of mammalian ribosomal subunits active in polypeptide synthesis in *Methods in Enzymology* (Moldave, K. and Grossman, L., eds.) Vol. 20, pp. 433–446, Academic Press, New York
15. Amils, R., Matthews, E.A., and Cantor, C.R. (1979) Reconstitution of 50S ribosomal subunits from *Escherichia coli* in *Methods in Enzymology* (Moldave, K. and Grossman, L., eds.) Vol. 59, pp. 449–461, Academic Press, New York
16. Ausubel, F.M., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A., and Struhl K. (1994) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York
17. Landt, O., Grunert, H.P., and Hahn, U. (1990) A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene* **96**, 125–128
18. Watanabe, K., Dansako, H., Asada, N., Sakai, M., and Funatsu, G. (1994) Effects of chemical modification of arginine residues outside the active site cleft of ricin A-chain on its RNA *N*-glycosidase activity for ribosomes. *Biosci. Biotechnol. Biochem.* **58**, 716–721
19. Bensadoun, A. and Weinstein, D. (1976) Assay of proteins in the presence of interfering materials. *Anal. Biochem.* **70**, 241–250
20. Sali, A. and Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815
21. Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291
22. Lüthy, R., Bowie, J.U., and Eisenberg, D. (1992) Assessment of protein models with three-dimensional profiles. *Nature* **356**, 83–85
23. Lin, Q., Chen, Z.C., Antoniw, J.F., and White, R.F. (1991) Isolation and characterization of cDNA clone encoding the anti-viral protein from *Phytolacca americana. Plant Mol. Biol.* **17**, 609–614
24. Hartley, M.R., Chaddock, J.A., and Bonness, M.S. (1996) The structure and function of ribosome-inactivating proteins. *Trends Plant Sci.* **1**, 254–260
25. Zoubenko, O., Hudak, K., and Tumer, N.E. (2000) A non-toxic pokeweed antiviral protein mutant inhibits pathogen infection via a novel salicylic acid-independent pathway. *Plant Mol. Biol.* **44**, 219–229
26. Stirpe, F., Barbieri, L., Gorini, P., Valbonesi, P., Bolognesi, A., and Polito, L. (1996) Activities associated with the presence of ribosome-inactivating proteins increase in senescent and stressed leaves. *FEBS Lett.* **382**, 309–312
27. Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**, 4876–4882
28. Kraulis, P. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structure. *J. Appl. Cryst.* **24**, 946–950
29. Rajamohan, F., Pugmire, M.J., Kurinov, I.V., and Uckun, F.M. (2000) Modeling and alanine scanning mutagenesis studies of recombinant pokeweed antiviral protein. *J. Biol. Chem.* **275**, 3382–3390
30. Watanabe, K. and Funatsu, G. (1986) Involvement of arginine residues in inhibition of protein synthesis by ricin A-chain. *FEBS Lett.* **204**, 219–222
31. Watanabe, K., Suemasu, Y., and Funatsu, G. (1989) Identification of lysine residues at or near active site of luffin-a, a ribosome-inactivating protein from seeds of *Luffa cylindrica. J. Biochem.* **106**, 977–981