Dissertation Submitted for
the Degree of Doctor of Philosophy

Structure and Function of Rosaceous S-RNase Associated with Gametophytic Self-Incompatibility

by
Takeshi Ishimizu

Division of Protein Chemistry
Institute for Protein Research
Osaka University

February, 1998

Dissertation Submitted for
the Degree of Doctor of Philosophy

Structure and Function of Rosaceous S-RNase Associated with Gametophytic Self-Incompatibility

by
Takeshi Ishimizu

Division of Protein Chemistry
Institute for Protein Research
Osaka University

February, 1998

Contents

Abbreviations

2D	two-dimensional
4-VP	4-vinylpyridine
API	*Achromobacter* protease I
Asn-C	Jack Bean asparaginyl endopeptidase
d_N	the numbers of nonsynonymous substitutions
d_S	the numbers of synonymous substitutions
GSI	gametophytic self-incompatibility
HPLC	high performance liquid chromatography
HV	hypervariable
LC	liquid chromatography
Mes	2-(*N*-morpholino)ethanesulfonic acid
MHC	major histocompatibility complex
MS	mass spectrometry
NEPHGE	non-equilibrium pH gradient electrophoresis
PA-	pyridylaminated
PAGE	polyacrylamide gel electrophoresis
PE-	*S*-pyridylethylated
PS region	region which positive selection may operate
PTH	phenylthiohydantoin
PVDF	polyvinylidene difluoride
RCM-	reduced and *S*-carboxymethylated
RPS	reversed-phase scale
SI	self-incompatibility
SLG	*S*-locus glycoprotein
sm	stylar-part mutant
SSI	sporophytic self-incompatibility
SRK	*S*-related kinase
V8	*Staphyrococcus aureus* V8 protease

General Introduction

Flowering plants produced the next generation by fertilization; fusion between the sperm nucleus in the tip of the pollen tube and the egg nucleus in the ovule of the pistil (Figure 1 (a)). Interestingly, many flowering plants have specific mechanisms to prevent inbreeding. The most common mechanism is a self-incompatibility (SI) system, defined as "the inability of a fertile hermaphrodite seed plant to produce zygotes after self-pollination" (de Nettancourt, 1977). Contemporary researchers agree with the plausible theory that self-incompatible reproduction (sexual reproduction) is more advantageous than self-compatible reproduction (asexual reproduction) because it creates genetic polymorphism, the raw material for evolution by natural selection, and results in the assembly of mutations that confer an advantage or in the clearance of those that do not (Lyons, 1997). Charles Darwin was the first to discuss this system and suggest its central significance during the evolution of flowering plants (Darwin, 1876). He was attracted to the system and stated that "self-sterility is one of the most surprising facts which I have ever observed" (Darwin, 1876). Like Darwin, many scientists—botanists, mathematical geneticists, molecular biologists, biochemists, and structure biologists— are interested in this phenomenon. The SI system has many mysteries that need solving and these offer attractive research projects in various fields of science.

SI is classified as one of two types, heteromorphic or homomorphic SI, on the basis of whether the incompatibility is associated with floral morphology. Heteromorphically self-incompatible species have short- and long-styled flowers, and self-pollination in each flower is physically difficult. In addition to the physical

exclusion of self-pollination, the pistil has a system which accepts only pollen grains from a different type of flower.

Homomorphic SI is distributed widely in a number of families; more than half the angiosperm species (de Nettancourt, 1977). In this SI type, generally there is no physical barrier to self-pollination, but the germination of self pollen or growth of the self pollen tube is inhibited in the stigma or style of the pistil. In most cases, this phenomenon is controlled by a single locus (*S*-locus) with multiple alleles (S_1, S_2, S_3, ..., S_k).

Homomorphic SI further is divided into two subgroups, sporophytic and gametophytic self-incompatibility (SSI and GSI), according to the *S*-phenotype of the pollen. The pollen *S*-phenotype is determined by the diploid genotype of the pollen parent in SSI or by the haploid genotype of pollen in GSI. In the SSI system, fertilization does not occur when either of the diploid *S*-alleles of the pollen parent matches either allele of the pistil. In the GSI system, shown in Figure 1 (b), fertilization does not occur when the haploid *S*-allele of the pollen matches either of the diploid *S*-alleles of the pistil.

SSI has been identified in six families, including the Cruciferae, Convolvulaceae, and Compositae (de Nettancourt, 1977). In the Cruciferae and Convolvulaceae, there are dominance relationships between the *S*-alleles that affect the compatible mating pairs. The pollen grain cells of these families are trinucleate, and the stigmas dry. Inhibition of incompatible pollen tube germination occurs at the surface of stigma. In *Brassica spp.* of the Cruciferae, glycoproteins that cosegregate with *S*-alleles and are expressed in the stigma of the pistil are called SLGs (*S*-locus glycoproteins) (Nasrallah *et al.*, 1985; Takayama *et al.*, 1987; Nasrallah *et al.*, 1987). A number of cDNAs that encode SLGs have been cloned, but their functions are unknown. SRKs (*S*-related kinases) that have a serine-threonine kinase domain, a transmembrane domain, and an extracellular domain with amino acid sequences similar to those of

SLGs, also cosegregate with the *S*-locus and are physically linked to the SLG genes (Stein *et al.*, 1991). Although much molecular biological research has been done, the complete picture of the molecular mechanism of this system is still not clear (for commentary see Nasrallah, 1997).

GSI has been reported in 15 families, including the Rosaceae, Solanaceae, Scrophulariaceae, Papaveraceae, and Poaceae (de Nettancourt, 1977). In gametophytic self-incompatible species (except the Papaveraceae, Poaceae, and Oenotheraceae in which the morphology of the SI resembles that of the SSI) the pollen grain, which is binucleate, germinates a pollen tube when it lands on the wet stigma surface of the pistil. This pollen tube then penetrates the stigma and grows between the longitudinal files of cells of the central transmitting tissue in the style toward the ovary. The incompatible pollen grain also germinates, and its pollen tube initially grows like the compatible pollen tube. In the upper style, however, the tube wall thickens, and callose deposition is seen near the tip. Growth of the tube slows, and the swollen tube tip sometimes bursts, resulting in no fertilization. This inhibition of pollen tube growth is, however, reversible because the pollen tube in the incompatible style in which growth is inhibited can grow again by grafting the incompatibly pollinated style onto the compatible style (Lush and Clarke 1997).

Molecular biological investigations of the GSI mechanism began in the early 1980s. The experimental findings obtained by 1994, when I started this study, are as follows.

In 1986, a glycoprotein (S_2-glycoprotein) that cosegregates with the S_2-allele of *Nicotiana alata* of the Solanaceae was isolated, and its cDNA cloned (Anderson *et al.*, 1986). cDNAs encoding S_3- and S_6-allele-specific glycoproteins also were obtained (Anderson *et al.*, 1989). Subsequently, about 30 cDNAs of S-glycoproteins have been cloned and sequenced from other solanaceous species, including *Petunia*, *Solanum*, and *Lycopersicon*. These S-

glycoprotens are composed of five conserved and two hypervariable regions (loerger *et al.*, 1991). These S-glycoproteins are allelic products of a single locus and a single copy gene in the genome (Anderson *et al.*, 1989). S-glycoprotein is expressed specifically in the stigmatic papillae and extracellular matrix that separates the files of the transmitting tract cells of the mature style, the way that guides the pollen tube (Cornish *et al.*, 1987). The increase in the expression of this protein during floral development parallels the acquisition and intensity of the SI character (Anderson *et al.*, 1986). Low level S-glycoprotein expression also occurs in the ovary and the developing pollen, but it is only 1 % that in the mature style (Dodds *et al.*, 1993; Clark and Sims, 1994).

The two conserved regions of the S-glycoproteins are homologous to the two regions essential for enzymatic activity of the fungal ribonuclease, RNase T_2. As it has been shown experimentally that S-glycoprotein actually has ribonuclease activity (McClure *et al.*, 1989; Kawata *et al.*, 1990), it has been called S-RNase. The S-glycoproteins of various solanaceous species also have ribonuclease activity (Singh *et al.*, 1991).

In 1994, stylar S-RNase was shown to be necessary and sufficient for SI interaction between the pistil tissue and pollen grain (Lee *et al.*, 1994; Murfett *et al.*, 1994) and ribonuclease activity of S-RNase to be essential for this interaction (Huang *et al.*, 1994). Many points, however, are unclear. What is the substrate of this ribonuclease? How does S-RNase come in contact with the pollen? What is the *S*-allele product expressed in pollen? What does the S-RNase interact with? Does it interact with the pollen S-product? How does S-RNase function in the discrimination of self/non-self pollen?

The molecular mechanism of the GSI of the Papaveraceae, however, is differ from that of the Solanaceae. Although low RNase activity has been detected in the poppy stigma, purified poppy S-glycoproteins do not have RNase activity (Franklin-Tong *et al.*, 1991; Foote *et al.*, 1994). This means that the molecular

mechanism of GSI is not universal and suggests that each family has its own GSI system.

The SI systems of members of the Rosaceae, such as *Pyrus* (pear) and *Malus* (apple), which also are controlled gametophytically by a single, multi-allelic locus, the *S*-locus (Crane, 1925; Kikuchi, 1929; Latimer, 1937), have been little studied. Some components that may be associated with SI have been isolated from *Prunus avium* (sweet cherry), but they have yet to be characterized (Mau *et al.*, 1982).

On the basis of the preceding information, I investigated the molecular mechanism of the GSI of *Pyrus pyrifolia* Nakai (Japanese pear) (Figure 2), a member of the Rosaceae. The advantages of studying the *P. pyrifolia* GSI are [1] Crossing experiments have shown that the SI in this species is controlled gametophytically by a single *S*-locus (Kikuchi 1929). Moreover, because seven *S*-alleles have been found so far in this species (Terami *et al.*, 1946), proteins that cosegregate with the *S*-locus can be unambiguously identified. [2] As *P. pyrifolia* belongs to the Rosaceae family, it is interesting to determine whether its system is similar to that in the Solanaceae or Papaveraceae, or is unique to the Rosaceae. [3] A self-compatible mutant of the cultivar 'Nijisseiki' (S_2S_4), 'Osa-nijisseiki', has been found (Furuta *et al.*, 1980). Its *S*-genotype was designated $S_2S_4^{sm}$ (sm: stylar-part mutant) in a crossing experiment (Sato, 1993). A comparison of the stylar proteins of 'Nijisseiki' and 'Osa-nijisseiki' therefore should provide direct evidence of the relationship between the stylar proteins and SI system.

Since 1991 a number of Japanese researchers have made molecular investigations of the GSI of *P. pyrifolia* (Hiratsuka, 1992; Nakanishi *et al.*, 1992; Sassa *et al.*, 1992). They could not identify the stylar protein encoded by the *S*-locus because all of them used a one-dimensional gel electrophoresis technique. Sassa *et al.* (1992), however, using isoelectric focusing with

subsequent staining for RNase activity, detected three stylar basic RNases correlated with *S*-alleles, indicative that the stylar protein encoded by the *S*-locus is RNase. Three *S*-glycoproteins were separated by two-dimensional gel electrophoresis (2D-PAGE) and indirectly identified as the basic RNases (Sassa *et al.*, 1993).

I did molecular level biological and biochemical studies to answer three questions: [1] Is the stylar S-protein of the Rosaceae an S-RNase? [2] What is the structure of this rosaceous S-protein? [3] Which region of the S-protein is involved in the SI interaction between the pistil and pollen? First, using 2D-PAGE, I identified the seven stylar proteins encoded by the *S*-locus of *P. pyrifolia* and determined their *N*-terminal amino acid sequences. Amino acid sequence analysis and enzymatic assay findings for these proteins showed that they are S-RNases similar to those present in solanaceous plants (Chapter I). The amino acid sequences of the proteins and the nucleotide sequences of their cDNAs next were analyzed to determine their primary structures. These structures were compared with those of other rosaceous S-RNases to ascertain their primary structural features (Chapter II). The structures of the post-translational modifications (disulfide bonds and *N*-glycans) of the *P. pyrifolia* S-RNases are reported in Chapters III and IV. To further characterize the primary structures of the rosaceous S-RNases and to be able to discuss the *S*-allele-specific recognition sites in rosaceous S-RNases, the numbers of synonymous (d_S) and nonsynonymous (d_N) substitutions in these proteins were calculated for each window (Chapter V). The relationship between the GSI system of the Rosaceae and the structural characteristics of the rosaceous S-RNase are discussed based on the structural information obtained. Finally, insight gained into the function of S-RNase in discriminating between self and non-self pollen in the RNase-based GSI system is discussed.

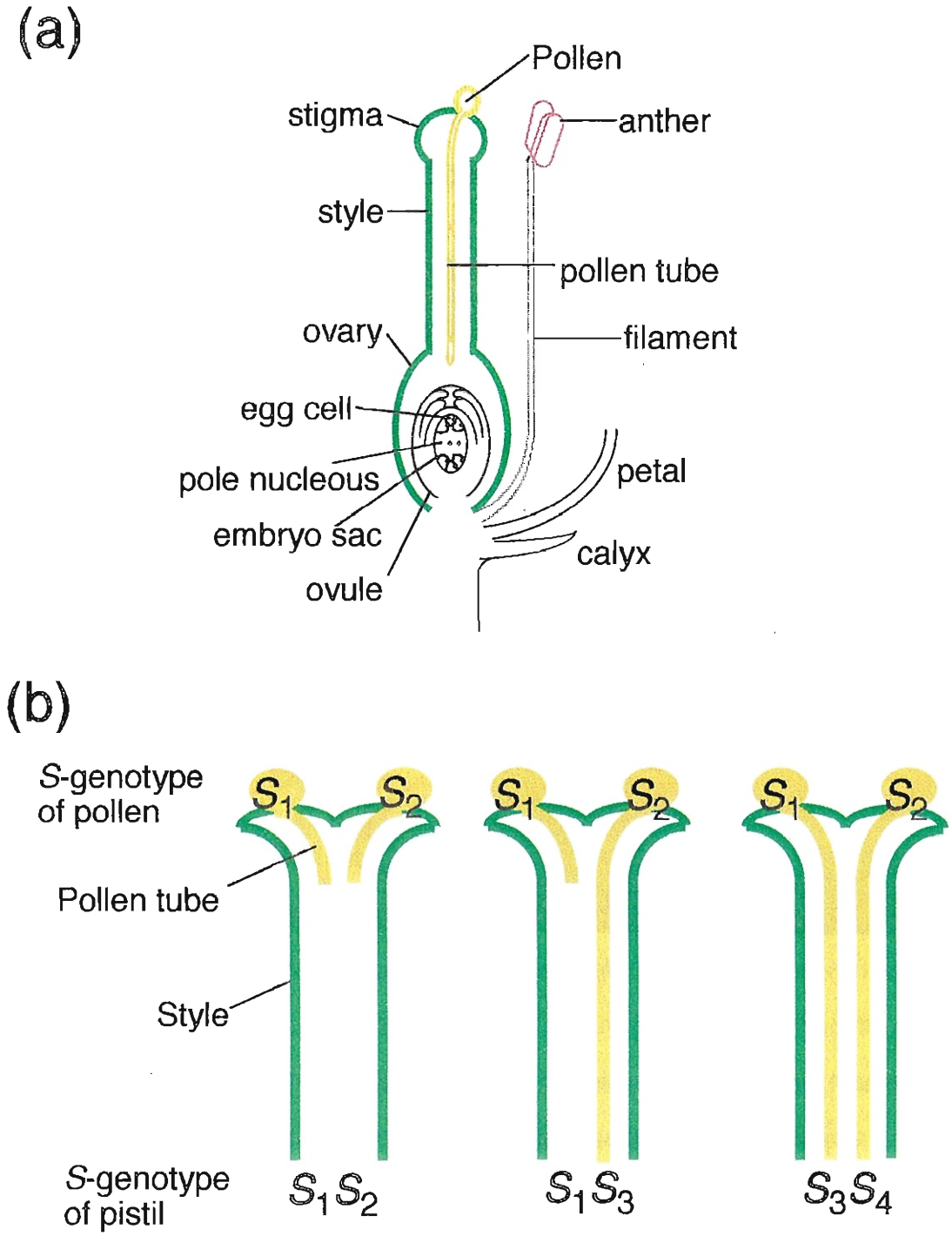


Figure 1. (a) **Structure of the flower.** When the sperm nucleus in the tip of the pollen tube reaches and fuses with the egg nucleus in the ovule of the pistil, fertilization occurs. (b) **Gametophytic self-incompatibility.** This system is controlled by the *S*-locus. When the haploid *S*-allele of the pollen matches one of the diploid *S*-alleles of the pistil, pollen tube growth is inhibited in the style, and no fertilization occurs.

Figure 2. ***Pyrus pyrifolia* (Japanese pear) flowers.** *P. pyrifolia* flowers once a year in Japan in the middle of April.

Chapter I
Identification of S-RNase expressed in the pistil of *P. pyrifolia*

Introduction

This chapter was designed to identify *S*-allelic products specific for all of the S_1- to S_7-alleles in the style of *P. pyrifolia* of the Rosaceae and to shed light on the mechanism by which self-incompatible 'Nijisseiki' mutated to self-compatible 'Osa-Nijisseiki'. To identify the *S*-allelic products in the style, I used 2D-PAGE combined with *N*-terminal and partial amino acid sequence analyses. Seven *S*-allele specific proteins from fourteen variants, including typical cultivars and offsprings of which the *S*-genotypes were elucidated by conventional crossing tests, were separated by the 2D-PAGE analysis and assigned to individual *S*-alleles. S-proteins purified by a series of chromatographies had ribonuclease activity, indicated that S-proteins were S-RNases. The relationship between the expression of S-RNases and self-compatibility of 'Osa-Nijisseiki' was also investigated. I discussed the characters of *P. pyrifolia* S-RNases and the similarities and differences of self-incompatibility system between the Rosaceae and the other families exhibited GSI.

Materials and Methods

Plant materials

The flowers of *P. pyrifolia* were collected in 1992 to 1994 at the Tottori Horticultural Experiment Station in Daiei, Tottori and the Fruit Tree Research Station, Ministry of Agriculture, Forestry and Fisheries of Japan in Tsukuba, Ibaraki. The cultivars of *P. pyrifolia* used in this investigation are as follows; 'Imamura-aki' (S_1S_6), 'Chojuro' (S_2S_3), 'Nijisseiki' (S_2S_4), 'Osa-Nijisseiki' ($S_2S_4^{sm}$) (sm; stylar-part mutant), 'Seigyoku' (S_3S_4), 'Kosui' (S_4S_5) and 'Okusankichi' (S_5S_7) at Tottori and 'Hayatama' (S_1S_2), S_2S_2 homozygote, S_3S_3 homozygote and offsprings ('267-4', '267-39' and '268-26') between 'Osa-Nijisseiki' (female) and 'Kosui' (male) at Tsukuba. The *S*-genotypes of these varieties have been assigned by crossing experiments (Terami *et al.*, 1946; Machida, 1982; , Sato, 1992; Sato, 1993). The styles (in this dissertation, the term style indicates the style with the stigma attached) of each flower were collected, rapidly frozen in liquid nitrogen and stored at -170°C until required.

2D-PAGE

Samples for electrophoresis were prepared as follows. The styles of each cultivar (from approximately 500 flowers) were ground in liquid nitrogen using a mortar and a pestle, extracted with lysis buffer (5 ml) (O'Farrell, 1975) containing 3 % Polyclar-AT (polyvinylpyrrolidone, GAF Chemicals Co.) and 1.5 % sodium ascorbate, and centrifuged at 25000 g for 30 min at 4 °C. The supernatant was stored at -80°C until use.

2D-PAGE was performed by non-equilibrium pH gradient electrophoresis (NEPHGE) in the first dimension and SDS-PAGE in the second dimension according to the method of O'Farrell *et al.* (1977) with slight modifications. The first-dimensional gel (2.5 x 130 mm) was composed of 9 M urea, 2 % Nonidet P-40, 4.73 % acrylamide, 0.27 % *N,N'*-methylene-bisacrylamide, 2 % Ampholine pH 3.5-10, 1 % Pharmalyte pH 8-10.5. Samples corresponding to approximately

5 flowers were applied to the cathode end of the gel and electrophoresed for 2100-2900 V·hr until the current in the gel was less than 100 μA. Then the gel was equilibrated with SDS sample buffer (Laemmli, 1970) for 1 hr and applied to 12 % SDS-polyacrylamide gel. After electrophoresis in the second dimension, proteins in the gel were detected by silver staining using 2D-Silver stain·II (Daiichi Pure Chemicals, Tokyo) or Coomassie brilliant blue R-250.

Preparation of S-RNases

The styles of 'Nijisseiki', 'Kosui' or 'Chojuro' (approximately 1000 flowers) were ground in liquid nitrogen as described previously and extracted with 50 mM MES / NaOH buffer, pH 6.5 (150 ml), containing 5 mM EDTA·2Na, 1.5 % sodium ascorbate and 3 % Polyclar-AT, for 30 min on ice. After centrifugation at 16000 g for 10 min at 4°C, the supernatant was collected and chromatographed on a CM-cellulose column (14 x 250 mm) equilibrated with 50 mM MES / NaOH buffer, pH 6.5, with a linear gradient of 0 to 0.5 M NaCl at 4 °C. The S-RNase fraction bearing RNase activity was further purified by reversed-phase HPLC on a Vydac C_4 column (4.6 x 250 mm) with a gradient of 2-propanol / acetonitrile (7:3, v / v) (0-80 %) in 0.1 % trifluoroacetic acid.

Amino acid sequence analyses

S-RNases separated by 2D-PAGE were electroblotted onto polyvinylidene difluoride (PVDF) membranes as described by Hirano and Watanabe (1990). Membrane-blotted *S*-RNase was excised, placed in a Blott cartridge (Perkin Elmer / Applied Biosystems) and sequenced with a gas-phase protein sequencer (470A Applied Biosystems). To analyze internal sequences, *S*-RNase was blotted on PVDF membrane, digested with API in 0.1 M Tris / HCl, pH 9.0, containing 10 % acetonitrile and 1 % reduced Triton X-100 (Fernandez *et al.*, 1994) for 24 hr at 37°C at a substrate / enzyme ratio of 50 / 1 (mol / mol), and

eluted with 0.1 % trifluoroacetic acid in 40 % acetonitrile by sonication for 5 min. Fragmented peptides were separated on a YMC-Pack MB-ODS column (2.1 x 50 mm) with a linear gradient of acetonitrile (0 - 40 %) in 0.1 % trifluoroacetic acid at a flow rate of 200 μl / min. Each peptide fragment thus obtained was sequenced with the protein sequencer.

Results

Identification of seven stylar S-proteins by 2D-PAGE and amino acid sequence analysis

Cultivars and offsprings of *P. pyrifolia* used in this experiment are listed with their genetically established *S*-alleles in Table I-1. Stylar proteins were extracted with lysis buffer in the presence of polyvinylpyrrolidone and sodium ascorbate. The style extract from each cultivar was submitted to 2D-PAGE in which NEPHGE, separating proteins of pI 4.5 to 10.5, and SDS-PAGE were used as the first and second dimensions, respectively. Figure I-1 shows a typical pattern of 2D-PAGE of the style extract of *P. pyrifolia*. Almost all proteins separated over the whole region of the gel were detected for all cultivars tested, but critical differences in the pattern were found in a particular region (hereafter called the S-protein zone) to which 30-32 kDa basic S-proteins migrated (Figures I-1 and 2). Positional comparison of these proteins for any of two cultivars sharing a given *S*-allele made it possible to assign tentatively each basic protein to one of the S_1- to S_5-alleles (Figure I-2). This assignment was then confirmed by analyzing the *N*-terminal amino acid sequence (Figure I-3). In this assignment, the S_6-protein of 'Imamura-aki' (S_1S_6) and the S_7-protein of 'Okusankichi' (S_5S_7) were assigned based on the position in the S-protein zone and the *N*-terminal

amino acid sequence since S_6- or S_7-allele is known to exist only in one of these cultivars. Seven stylar S-proteins separated on a gel were eventually assigned to individual *S*-genotypes. A composite of the locations of these assigned proteins is presented in Figure I-4. Detailed procedures for the identification of S-proteins are described below.

a) S_1-protein—'Imamura-aki' (S_1S_6) gave two separate protein spots in the S-protein zone. One was a 30.5 kDa protein which migrated slower than the other one in NEPHGE. The same 30.5 kDa protein was detected for 'Hayatama' (S_1S_2), and was assigned as S_1-protein. Its *N*-terminal amino acid sequence was determined as YDYFQFTQQYxPAVxN (x denotes an unidentified residue). The YDYFQFTQQYWPAV sequence has been reported for the *N*-terminus of S_1-RNase (Sassa *et al.*, 1993).

b) S_2-protein—In the S-protein zone, 'Nijisseiki' (S_2S_4) gave four protein spots which were divided into two groups, each composed of two proteins. One group with a molecular mass of 32 kDa was found for 'Hayatama' (S_1S_2), 'Osa-Nijisseiki' ($S_2S_4^{sm}$), an S_2S_2 homozygote and two offsprings ('267-39' (S_2S_5) and '268-26' (S_2S_5)) derived from 'Osa-Nijisseiki', and was assigned as S_2-protein. The two components of S_2-protein were named S_2a- and S_2b-proteins (low and high mobility on SDS-PAGE, respectively). Only S_2b-protein was detected in other S_2-allelic cultivars such as 'Chojuro' (S_2S_3) and 'Kikusui' (S_2S_4). The *N*-terminal amino acid sequence analysis of PVDF membrane-blotted proteins revealed that S_2a- and S_2b-proteins from 'Nijisseiki' and S_2b-protein from 'Hayatama' shared the same sequence, ARYDYFQFTQQYQxAF. For most of the S_2-allelic cultivars tested, S_2b-protein was the major component.

c) S_3-protein—'Chojuro' (S_2S_3) gave three proteins in the S-protein zone. Of the three, the protein which migrated in the first dimension on gel electrophoresis had been assigned as S_2b-protein as described above. The remaining two proteins were also detected for 'Seigyoku' (S_3S_4) and the S_3S_3 homozygote.

The amounts of these two proteins estimated from the intensity of silver staining were roughly equal and they were named S_3a- (32 kDa) and S_3b- (31 kDa) proteins on the same basis as mentioned earlier. The *N*-terminal amino acid sequences of both S_3a- and S_3b-proteins were the same, being YDYFQFTQQYxLAVxN. The mobility of S_3b-protein on NEPHGE was the smallest among the seven *S*-proteins tested.

d) S_4-protein—For 'Nijisseiki' (S_2S_4), two protein spots with similar intensities on silver staining migrated more slowly than S_2a- and S_2b-proteins in the first dimension. These two slowly migrating proteins, which were also found for 'Kikusui' (S_2S_4), were assigned as S_4a- (31 kDa) and S_4b- (30.5 kDa) proteins. 'Seigyoku' (S_3S_4) and 'Kosui' (S_4S_5) gave S_4b- protein only. The *N*-terminal FDYFQFTQQYQPAVxN sequence was found for either S_4a-protein or S_4b-protein from 'Nijisseiki'. All residues except C-terminal Asn have been determined for S_4-RNase (Sassa *et al.*, 1993).

e) S_5-protein—In addition to S_4b-protein, two slower-migrating proteins were clearly detected by 2D-PAGE for 'Kosui' (S_4S_5) and were assigned as S_5a- (32 kDa) and S_5b- (31kDa) proteins. S_5a-protein was a major component of Kosui, and the same compositional feature was observed for offspring '267-39' (S_2S_5). 'Okusankichi' (S_5S_7) shared S_5a-protein with Kosui and offspring '268-26' (S_2S_5) gave S_5b-protein as a major component. The *N*-terminal amino acid sequence, YDYFQFTQQYQLAVxN, was found for both S_5a- and S_5b-proteins from Kosui. All residues except the eleventh were identical with those of S_3-protein as described above or with those of S_5-RNase (Sassa *et al.*, 1993).

f) S_6-protein—S_6-protein, a 31 kDa protein, was detected as a single spot together with S_1-protein identified for 'Imamura-aki' (S_1S_6) and had the highest mobility in the first direction among the seven S-proteins identified by 2D-PAGE. The spot of S_6-protein on the gel was not superimposable on that of any of the S-proteins separated from other cultivars. The *N*-terminal amino acid sequence

was found to be YNYFQFTQQYxPAVxN. The sequence contains all of the nine conserved residues (underlined), though Asn2 was distinct from the Asp counterpart observed for S_1- to S_5-proteins and S_7-protein. From these results, the 31 kDa protein was assigned as S_6-protein.

g) S_7-protein—'Okusankichi' (S_5S_7) gave a protein spot with a slightly faster mobility than S_5a-protein in the S-protein zone on SDS-PAGE. This spot was distinct from all other S-proteins that have ever been identified, and was assigned as S_7-protein. The sequence YDYFQFTQQYxPAV was found for its *N*-terminus, which is the same as the *N*-terminal sequence of S_1-protein.

S-Proteins are S-RNases

S-RNases, which are RNases cosegregated with S-locus, have been successfully purified by a series of chromatographic steps monitored with assay for ribonuclease activity from the style of *P. pyrifolia* in our laboratory. To see whether S_4-protein is active S_4-RNase, chromatographically purified S_4-RNase from 'Nijisseiki' or 'Kosui' was mixed with the style extract of the former cultivar and submitted to 2D-PAGE. The S_4-RNase, though stained as a diffuse spot, co-migrated with S_4a- and S_4b-proteins on the gel, thus identifying the S_4-protein as S_4-RNase (Figure I-5, B). When S_3-RNase isolated from Chojuro was electrophoresed by the same procedure, the enzyme appeared as a sole additional protein at the position to which the S_3-protein migrates on the gel, leading to the assignment of S_3-protein as S_3-RNase (Figure I-5, C).

To acquire crucial evidence for the assignment of S_2- and S_4-proteins as S_2- and S_4-RNases, internal amino acid sequences of S_2b- and S_4b-proteins from 'Nijisseiki' were analyzed. Briefly, the protein electroblotted onto PVDF membrane after 2D-PAGE separation was digested with API on the membrane. Fragment peptides were separated by reversed-phase HPLC and sequenced. As a result, 81 and 127 residues, all comprising independent peptides, were

sequenced for S_2- and S_4-RNases, respectively. In the sequences determined by these analyses, two short stretches of the peptide chains including counterparts corresponding to two essential histidine residues of RNase T_2 (Kawata *et al.*, 1990) were found (Figure I-6). From this and the previous results, it was concluded that S_2- and S_4-proteins purified by 2D-PAGE are members of the RNase T_2 family and they were renamed S_2- and S_4-RNases, respectively. Thereafter, S-proteins separated by the same method as that used for these two S-RNases will be called S-RNases.

S_4-RNase is not expressed in self-compatible 'Osa-Nijisseiki'

When the style extract of 'Osa-Nijisseiki' ($S_2S_4^{sm}$) was analyzed by 2D-PAGE, S_2-RNase was clearly detected as S_2a- and S_2b-proteins. However, no proteins corresponding to S_4a- and S_4b-proteins were detected in the S-protein zone (Figure I-7), leading to the notion that no detectable S_4-RNase exists in the style of Osa-Nijisseiki. The absence of S_4-RNase was also observed for a self-compatible offspring, 267-4 ($S_4^{sm}S_5$), derived from crossing Osa-'Nijisseiki' (female) and 'Kosui' (male). From this result, a close relationship of self-compatibility of 'Osa-Nijisseiki' with failure of the synthesis of S_4-RNase in the style is suggested.

S-RNases are developmentally expressed

The appearance of *S*-RNases in the style of 'Nijisseiki' was followed by 2D-PAGE during flower development. Figure I-8 depicts the developmental appearance of S_2- and S_4-RNases in the S-protein zone at five stages from green bud, pink bud, white bud, balloon and mature flower. The two RNases were hardly detectable at the green bud and pink bud stages and were clearly detected at subsequent stages up to anthesis, indicating that these S-RNases are developmentally expressed during the process of flower maturation.

Discussion

The present 2D-PAGE results showed that seven *S*-allele specific proteins, S-RNases, of *P. pyrifolia* are separated without superimposition (Figures I-1, 2, and 4). These S-RNases were assigned to individual *S*-alleles by positional comparison with those from cultivars bearing common *S*-genotypes and in most cases by comparison of their *N*-terminal amino acid sequences. A similar experiment has been reported by Sassa *et al.* (1993), who analyzed the *N*-terminal amino acid sequences of three S-glycoproteins, which were identified as S_1-, S_4- and S_5-RNases. However, for some reason S_2-RNase has never been detected by 2D-PAGE in cultivars such as 'Nijisseiki' (S_2S_4), 'Hayatama' (S_1S_2), 'Doitsu' (S_1S_2) and 'Osa-Nijisseiki' ($S_2S_4^{sm}$). Since this procedures and conditions for gel electrophoresis are not much different from those used for the earlier experiments, successful detection of S_2-RNase, a structurally distinct S-RNase molecule among *P. pyrifolia* S-RNases, on the gel is probably due to direct extraction with lysis buffer of the S-RNases from the *P. pyrifolia* style. In any event, the present electrophoretical separation of seven S-RNases was successful, since all detectable S-RNases exclusively migrated to the S-protein zone. This method should also be applicable for the determination of *S*-genotypes of new cultivars which would be derived by crossing from known cultivars of *P. pyrifolia*.

It is important that the present method apparently separates subcomponents of S-RNases, which are often detected as a double spot. In these cases, determination of the identical *N*-terminal amino acid sequence for the two proteins is necessary but not sufficient to establish the identity of their amino acid

sequences, since S_3-RNase and S_5-RNase electrophoretically behave differently but have the same sequence. Eventually, I compared the peptide map of the API digest of each subcomponent on a reversed-phase column under acidic conditions. For instance, the peptide maps of S_3a-RNase and S_3b-RNase were indistinguishable (data not shown). Since *N*-glycosylated peptides generally have very similar retention volumes on a reversed-phase column under acidic conditions, this observation supports the idea that the two subcomponent proteins which migrated at different positions have the same amino acid sequence. In fact, S_4-RNase, a mixture of S_4a- and S_4b-RNases, was detected as a single peak on a reversed-phase column, although two or more molecular ions were actually detected by mass spectrometry for the S_4-RNase purified chromatographically (data not shown). It is, therefore, likely that microheterogeneity of sugar moieties in an S-RNase is responsible for the occurrence of subcomponents separable by 2D-PAGE. The presence of a double spot for a given S-RNase of potato has also been reported (Thompson *et al.*, 1991).

The *N*-terminal amino acid sequence of S-RNases has a structural motif characteristic of *P. pyrifolia* proteins; $^{Y}/_{F}{}^{D}/_{N}$YFQFTQQYxx$A^{V}/_{F}$ is conserved for all seven S-RNases identified (Figure I-3). It is noteworthy that the *N*-terminal sequence of S-RNases from *M. domestica* is YDYFQFTQQYQPAV (Sassa *et al.*, 1994), which is highly homologous to that of *P. pyrifolia* S-RNases. These motifs are structurally distinct from those of solanaceous S-RNases. In addition, *P. pyrifolia* S_2-RNase is unique in that the Ala-Arg sequence is attached to the *N*-terminus of the above motif. The nucleotide sequence analysis of cDNA encoding S_2-RNase revealed that the protein is synthesized as a precursor bearing a signal peptide in which the Ala-Ala-Arg sequence is followed by Tyr1 (Norioka *et al.*, 1995; Norioka *et al.*, 1996). This indicates that a signal peptidase in the *P. pyrifolia* style can cleave the Ala-Ala bond but can not cleave the Arg-

Tyr bond in pro-S_2-RNase. The fact that the Gly-Phe and the Gly-Tyr bonds of the precursors of *P. pyrifolia* S_4-RNase (Norioka *et al.*, 1995; Norioka *et al.*, 1996) and *M. domestica* S_2-RNase (Sassa *et al.*, 1993), respectively, are sensitive to signal peptidase indicates that this peptidase favors non-polar side chains at the site of action in the propeptide-mature protein junction. These sites conform to the (-3, -1)-rule (von Heijine, 1986). This leads to the notion that, as in the case of *P. pyrifolia* S_2-RNase, the *N*-terminus of mature *Malus x domestica* S_3-RNase is Val-Lys-Phe and is not Phe since Ala-Val-Lys-Phe is located at positions -3 to +1 (Broothaerts *et al.*, 1995).

Expression of two S-RNases in the style of 'Nijisseiki' increased significantly between the pink bud stage and the balloon stage. These stages appear 1 to 3 days prior to anthesis and roughly correspond to the time of acquiring and enforcing SI (Hiratsuka *et al.*, 1985). Since, under the present conditions, no other detectable protein seems to be synchronized with acquisition of SI, our observation supports the idea that S-RNase is the key protein in the system segregating self pollen from non-self pollens in the pear style.

According to the results of genetic analyses, 'Osa-Nijisseiki', a self-compatible mutant of Nijisseiki, is a heterozygous stylar-part mutant ($S_2S_4^{sm}$) in which the S_4 gene of the style is exclusively mutated (Sato, 1993). Throughout our experiments, neither S_4a-protein nor S_4b-protein was detected in the style, in contrast to S_2-RNase which is always detectable in both 'Nijisseiki' and 'Osa-Nijisseiki'. No exception has been found in any examined style sample collected from any of the trees of 'Osa-Nijisseiki' at different years. S_4-RNase was also absent in the self-compatible offspring derived from 'Osa-Nijisseiki' (female) and 'Kosui' (male), suggesting that the potency for accepting self pollen of the former cultivar is genetically transferred to its offsprings. This result is consistent with the fact that the anti-sense targeted suppression of the expression of a given S-RNase abolishes the appearance of self-incompatibility in *Petunia inflata* of the

Solanaceae (Lee *et al.*, 1994). It is interesting to consider possible reasons why 'Nijisseiki' has gained self-compatibility by mutation. A possibility is that the S_4-RNase gene is partially or completely deleted in 'Osa-Nijisseiki'. In fact, Sassa *et al.* (1997) showed the suggestive data that S_4^{sm}-allele lacked S-RNase gene from the PCR analysis for 'Osa-Nijisseiki'. From the PCR analysis for the S_4^{sm} homozygote, it was proved that S_4-RNase gene was deleted in S_4^{sm}-allele (Norioka, N. *et al.*, unpublished data).

P. pyrifolia S-RNases exclusively migrated to the S-protein zone, to which ca. 30 kDa basic proteins move. Solanaceous S-RNases also migrated to the same zone (Mau *et al.*, 1986). Both S-RNases are of the RNase T_2 type. Moreover, the absence of style S_4-RNase is associated with self-compatibility in 'Osa-Nijisseiki'. These results support the idea that S-RNase is associated with GSI in rosaceous plants as well as in solanaceous plants.

Recently, SI genes were cloned in *Antirrhinum hispanicum* in the Scrophulariaceae and *Phalaris coerulescens* in the Poaceae of which SI were gametophytically controlled by *S*-locus and two unlinked *S*- and *Z*-loci, respectively. Stylar S-RNase are responsible for expression of SI character of in the Scrophulariaceae (Xue *et al.*, 1996). In the Poaceae, an S-protein expressed in pollen showed thioredoxin activity, and thioredoxin activity significantly reduced in S-protein from a self-fertile mutant (Li *et al.*, 1994; Li *et al.*, 1996). The molecular mechanism of GSI in the Rosaceae resembles those of the Solanaceae and the Scrophulariaceae. It is interesting to examine whether the origin of SI is common in these three families.

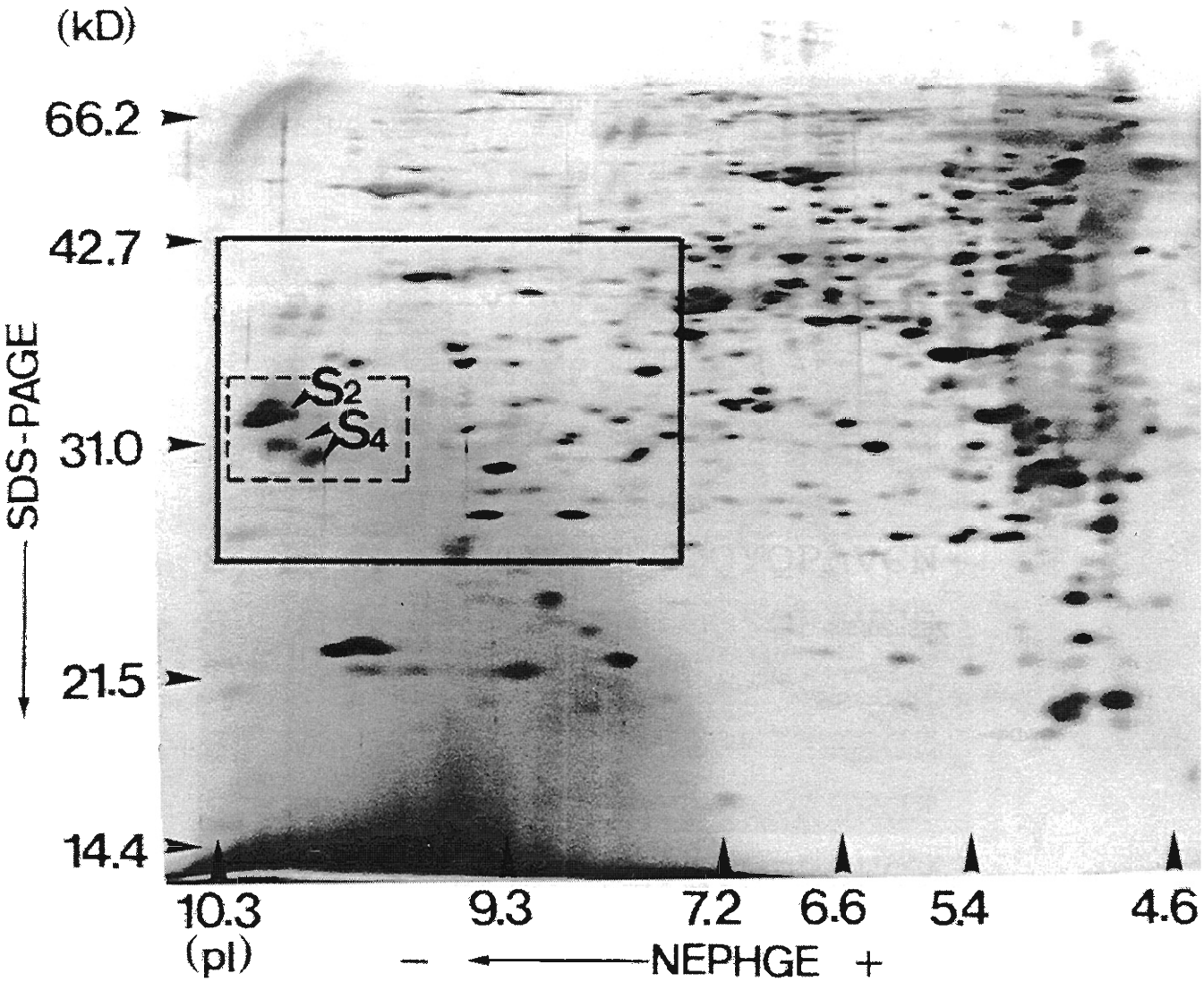


Figure I-1. **Pear stylar proteins separated by 2D gel electrophoresis.** Stylar proteins from 'Nijisseiki' (S_2S_4) were separated by 2D gel electrophoresis (NEPHGE / SDS-PAGE) and detected by silver staining. Numbers in the vertical column indicate molecular masses of standard proteins and those in the horizontal column denote their pI's. The S-protein zone and the zone presented in Figures I-2, 4, 5, 7 and 8 are boxed by broken and solid lines, respectively.

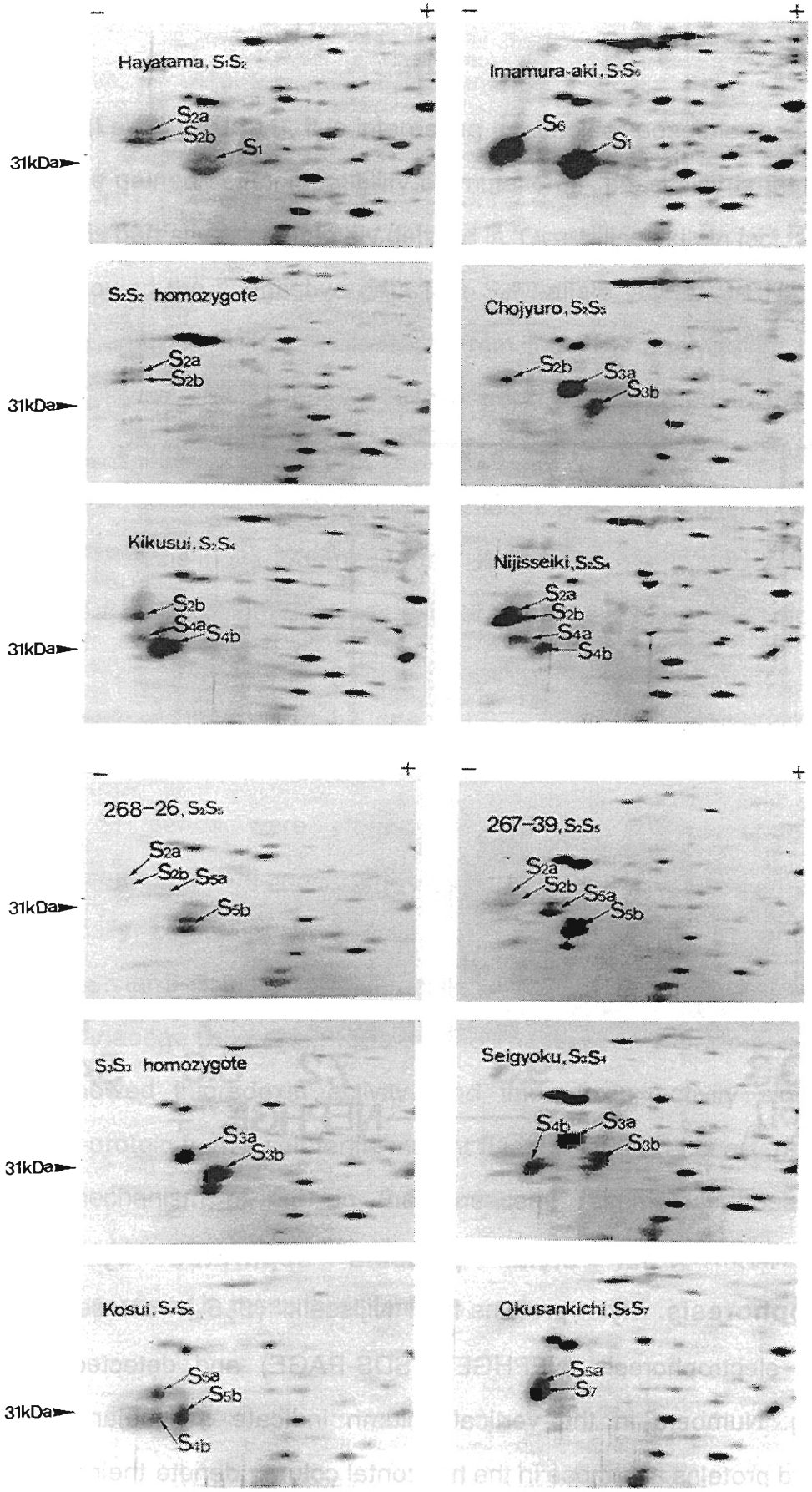


Figure I-2. **Separation of S-proteins from twelve varieties.** Each panel shows the separation of proteins in the S-protein zone. Assigned S-protein spots are indicated by arrows with *S*-alleles.

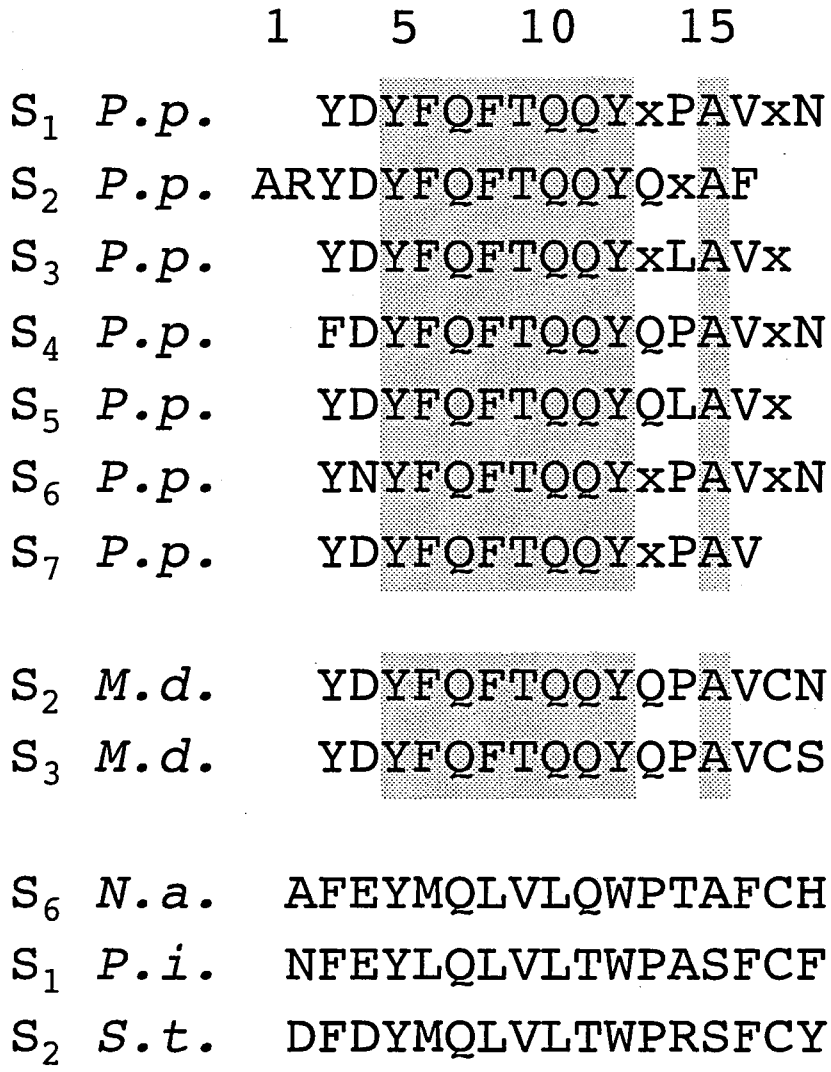


Figure I-3. **Comparison of *N*-terminal amino acid sequences of S-RNases from *P. pyrifolia*, *M. domestica* and the solanaceous plant.** S1 to S7 *P.p.*, S2 and S3 *M.d.*, S6 *N.a.*, S1 *P.i.* and S2 *S.t.* denote S1- to S7-RNases from *P. pyrifolia*, S2- and S3-RNases from *M. domestica* (Broothaerts et al., 1995), S6-RNase from *Nicotiana alata* (Anderson et al., 1989), S1-RNase from *Petunia inflata* (Ai et al., 1990) and S2-RNase from *Solanum tuberosum* (Kirch et al., 1989), respectively. Residues conserved in rosaceous S-RNases are shadowed. x indicates an unidentified residue.

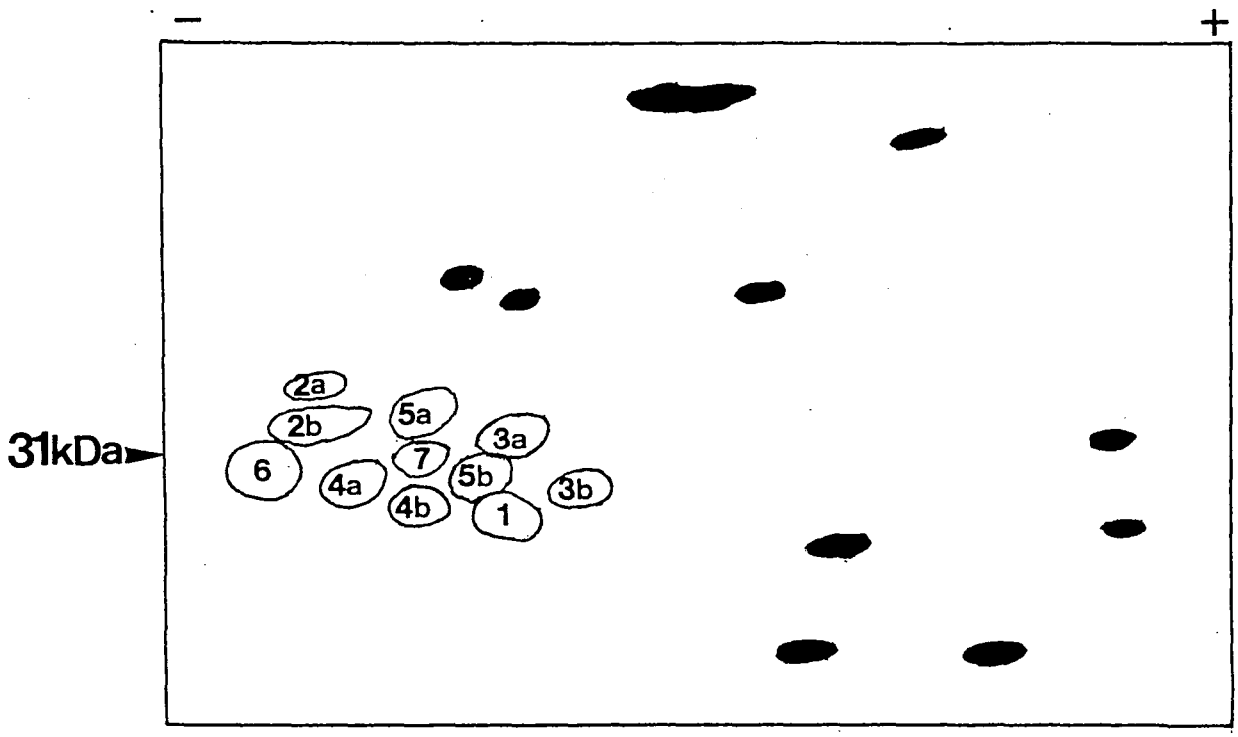


Figure I-4. **A composite for the location of seven S-RNases in the S-protein zone.** A composite panel was prepared by superimposing all panels shown in Figure I-2. Major proteins are shown by solid black. Open circles and numbers show the locations of S-proteins and their identified *S*-alleles, respectively.

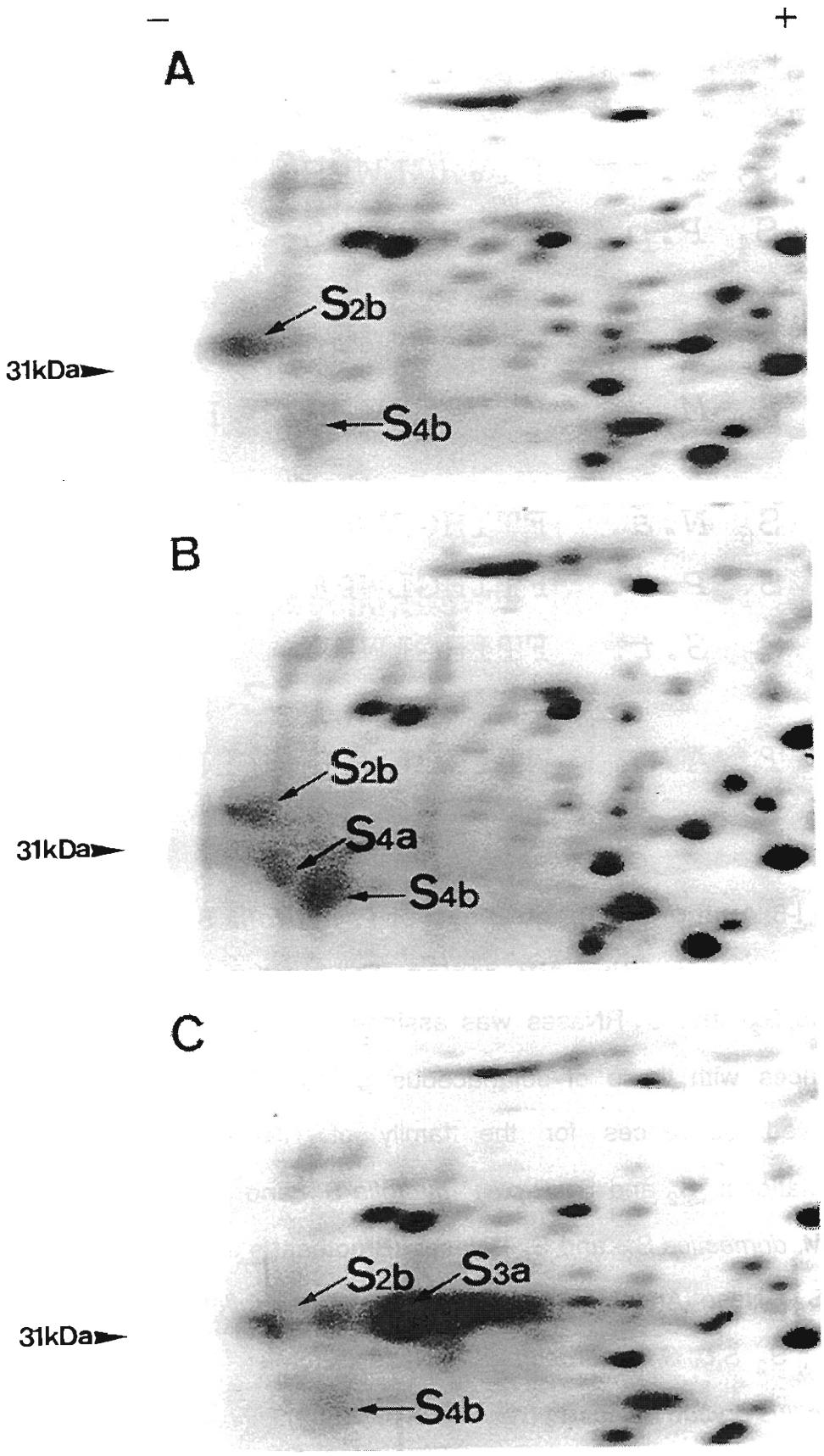


Figure I-5. **Co-migration of purified S_4- and S_3-RNases in the style extract of Nijisseiki to the S-protein zone.** (A), S_2- and S_4-RNases from the style extract of 'Nijisseiki' (S_2S_4) ; (B), (A) + purified S_4-RNase ; (C), (A) + purified S_3-RNase.

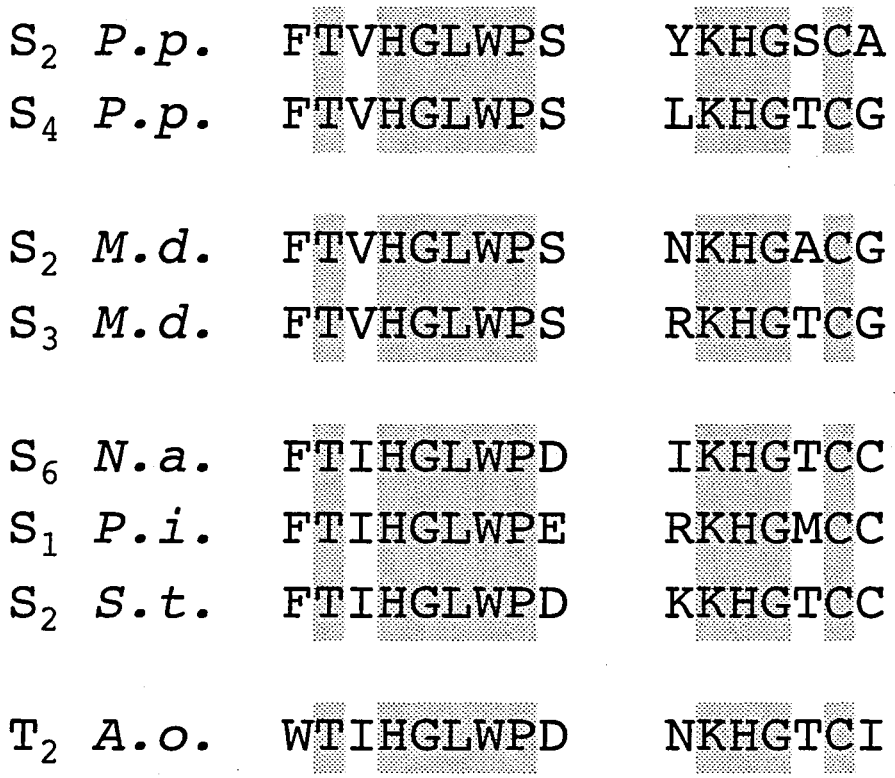


Figure I-6. **Amino acid sequences of *P. pyrifolia* S_2- and S_4-RNases in the putative active site region.** The putative active site region of *P. pyrifolia* S_2- and S_4-RNases was assigned by comparing the amino acid sequences with those of solanaceous S-RNases and RNase T_2. Totally conserved sequences for the family of RNase T_2 are shadowed. Abbreviations: S_2 and S_4 *P. p.*, *P. pyrifolia* S_2- and S_4-RNases ; S_2 and S_3 *M.d.*, *M. domestica* S_2- and S_3-RNases (Broothaerts et al., 1995) ; S_6 *N.a.*, *N. alata* S_6-RNase (Anderson et al., 1989) ; S_1 *P.i.*, *P. inflata* S_1-RNase (Ai et al., 1990) ; S_2 *S.t.*, *S. tuberosum* S_2-RNase (Kirch et al., 1989) and T_2 *A.o.*, *Aspergillus oryzae* RNase T_2 (Kawata et al., 1989).

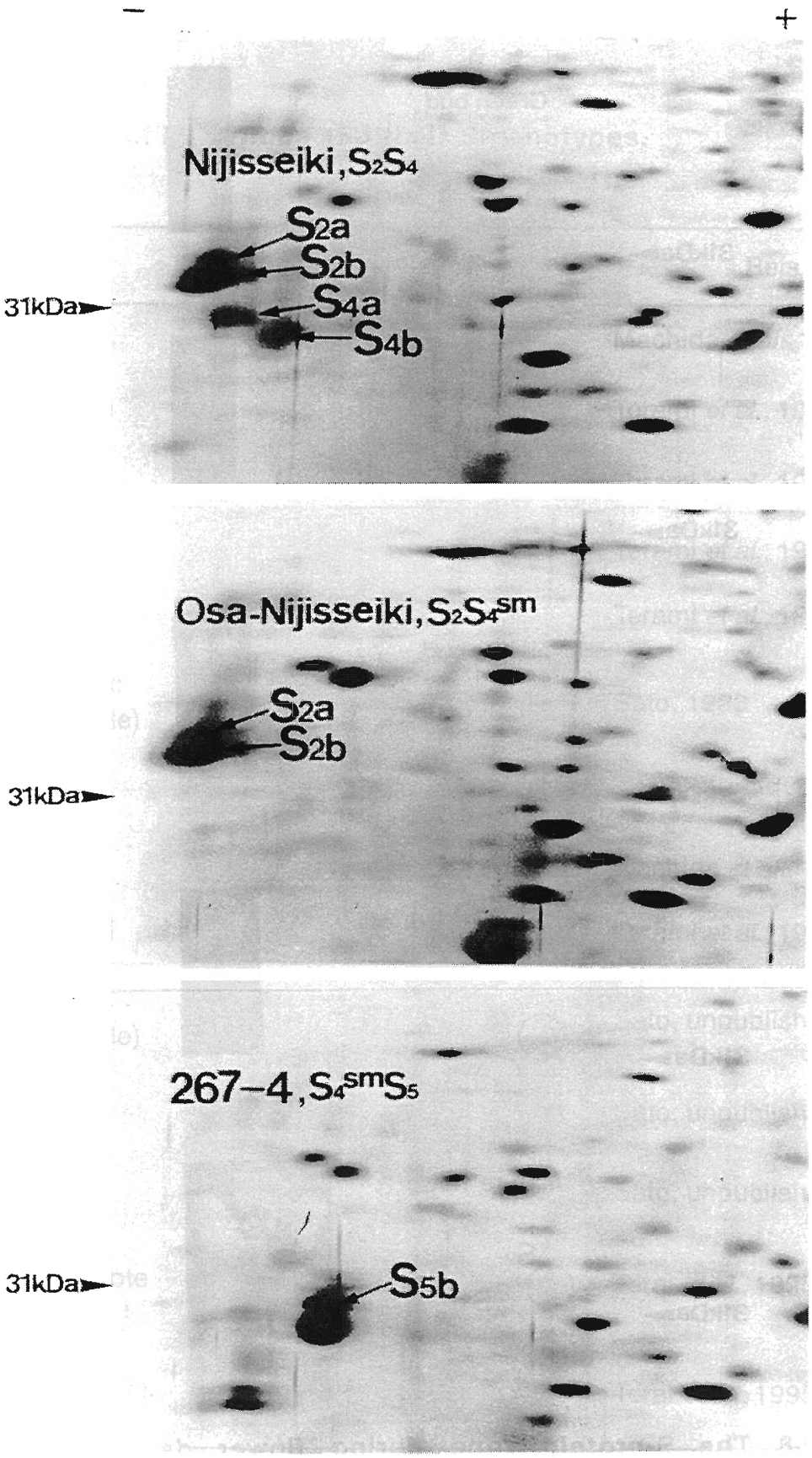


Figure I-7. **The S-protein zone of 'Osa-Nijisseiki' and its self-compatible offspring.** (A) 'Nijisseiki' (S_2S_4), (B) 'Osa-Nijisseiki' ($S_2S_4^{sm}$) and (C) an offspring, '267-4' ($S_4^{sm}S_5$).

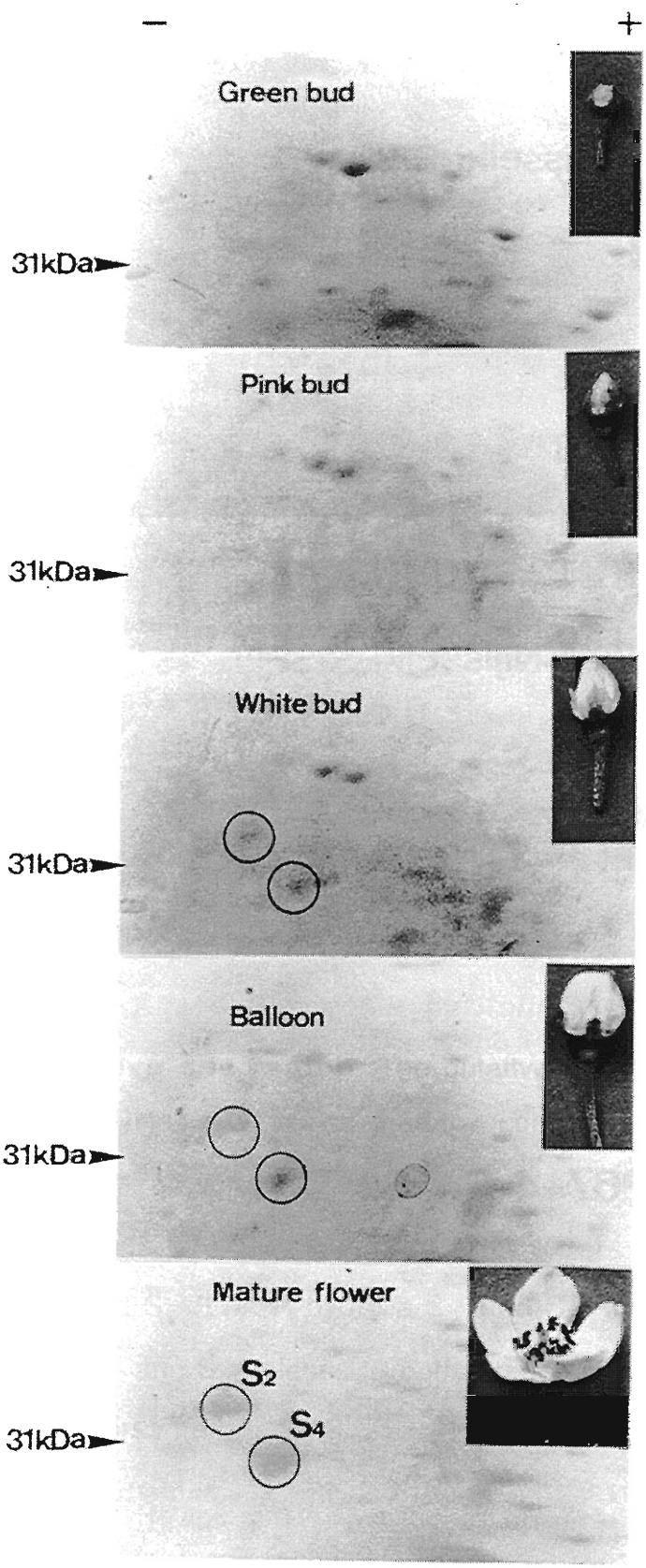


Figure I-8. **The S-protein zone during flower development of 'Nijisseiki'.** Stylar proteins from each stage of 'Nijisseiki' were separated by 2D gel electrophoresis and detected by staining with Coomassie brilliant blue R-250. The S-protein zone and the picture of the flower at each stage are shown. The spots of S_2- and S_4-RNases are marked with a circle.

Table I-1. **Cultivars of *P. pyrifolia* and their *S*-genotypes.**

Cultivar	*S*-genotype	Refs.
Hayatama	S_1S_2	Machida *et al.*, 1972
Imamura-aki	S_1S_6	Terami *et al.*, 1946
Chojuro	S_2S_3	Terami *et al.*, 1946
Kikusui	S_2S_4	Terami *et al.*, 1946
Nijisseiki	S_2S_4	Terami *et al.*, 1946
Osa-Nijisseiki (self-compatible)	$S_2S_4^{sm}$	Sato, 1993
Seigyoku	S_3S_4	Sato, 1992
Kosui	S_4S_5	Machida *et al.*, 1972
Okusankichi	S_5S_7	Terami *et al.*, 1946
267-4 (self-compatible)	$S_4^{sm}S_5$	Sato, unpublished result
267-39	S_2S_5	Sato, unpublished result
268-26	S_2S_5	Sato, unpublished result
S_2S_2 homozygote	S_2S_2	Terai *et al.*, 1995
S_3S_3 homozygote	S_3S_3	Terai *et al.*, 1995

Chapter II
Primary structural features of rosaceous S-RNases

Introduction

In Chapter I, I identified seven S-RNases which were expressed in the pistil, encoded by *S*-locus and associated with GSI in *P. pyrifolia*. Partial amino acid sequences of these S-RNases were determined. Based on these sequence information, Norioka and Ohnishi *et al.* (1995) in our laboratory cloned cDNAs of *P. pyrifolia* S_2- and S_4-RNase from a stylar cDNA library. S-RNases of *M. domestica* also identified (Sassa *et al.*, 1994; Broothearts *et al.*, 1995). cDNAs of *M. domestica* S_2- and S_3-RNases were cloned by RT-PCR on stylar RNA (Broothaerts *et al.*, 1995; Janssens *et al.*, 1995; Sassa *et al.*, 1996). Nucleotide sequence analyses of these clones showed that they had two consensus sequence motifs required for RNase activity in the RNase T_2 family, as well as solanaceous S-RNases, and they formed a subgroup distinct from solanaceous S-RNases in the neighbor-joining phylogenetic tree (Xue *et al.*, 1996).

But the function and the role of each region for the discrimination process of pollen *S*-alleles have not been clarified. In the progress of these studies, S-RNase was proved to be directly associated with SI by transformation experiments with sense and antisense S-RNase genes in solanaceous plants (Lee *et al.*, 1994; Murfett *et al.*, 1994). But it remains to be clear how S-RNase discriminates between self and non-self pollen.

To survey the primary structural features of rosaceous S-RNase and to discuss the *S*-allele-specific recognition site in the S-RNase, I determined the primary structures of *P. pyrifolia* S_1-, S_3-, S_5-, S_6-, and S_7-RNases and characterized the primary structures of rosaceous S-RNases using the alignment of amino acid sequences, pairwise comparison, and neighbor-joining phylogenetic tree. The hypervariable (HV) region in which many amino acid substitutions occurred and a highly homologous pair of *P. pyrifolia* S-RNases were detected. The recognition process in the SI reaction by S-RNase is discussed on the basis of these findings. The difference between the rosaceous and solanaceous S-RNases gene groups is also discussed.

Materials and Methods

Plant material

The cultivars of *P. pyrifolia*, 'Hosui' (S_3S_5), Kosui(S_4S_5), 'Imamuraaki' (S_1S_6) and 'Okusankichi' (S_5S_7), were obtained from the Tottori Horticultural Experiment Station, Daiei, Tottori, Japan and the Fruit Tree Research Station, Ministry of Agriculture, Foresty and Fisheries of Japan, Tsukuba, Ibaraki, Japan in 1995 and 1996. The *S*-genotypes of these cultivars have been assigned by crossing experiments (Terami *et al.*, 1946) or two-dimensional gel electrophoresis of the stylar proteins (Ishimizu *et al.*, 1998a). Styles with the stigma of each flower were collected, rapidly frozen in liquid nitrogen, then stored at -170°C until required.

Isolation of S-RNase

Five S-RNases (S_1-, S_3-, S_5-, S_6- and S_7-RNases) were purified by CM-cellulose column chromatography and reversed-phase HPLC (Chapter I, Ishimizu *et al.*, 1996a).

Protein sequence analysis

The S-RNases were reduced and *S*-carboxymethylated by the method of Crestfield *et al.* (1963). The carboxymethylated proteins were digested with *Achromobacter* protease I (API) in 20 mM Tris-HCl buffer, pH 9.0, at 37 °C for 6 hr at the molar substrate-to-enzyme ratio of 200 to 1, or with *Staphyrococcus aureus* V8 protease (V8) in 20 mM Tris-HCl buffer, pH 9.0, at 37 °C for 16 hr at the molar ratio of 100 to 1, or Jack Bean asparaginyl endopeptidase (Asn-C) (Takara) according to the manufacturer's instructions.

Fragmented peptides were separated by reversed-phase HPLC in a μBondasphere 5mm C_{18} 300Å column (Waters 3.9 x 150 mm). Sequence analyses of the carboxymethylated proteins and separated peptides were done with an Applied Biosystems 470A protein sequencer or a Hewlett Packard G1005A protein sequencer. Amino acid analyses of the undigested proteins and separated peptides were done with a Hitachi L-8500S amino acid analyzer after hydrolysis for 24 hr at 110 °C in evacuated tubes with twice-distilled 5.7 M hydrochloride containing 0.2% phenol.

Preparation of cDNA

Total RNA was isolated by the phenol-SDS procedure in combination with 0.3M sodium acetate precipitation and centrifugation in cesium chloride solution. The styles (white bud stage) were ground in liquid nitrogen using a mortar and a pestle. RNA was extracted with extraction buffer (200 mM Tris-HCl, pH 9.0, 100 mM NaCl, 10 mM EDTA, 0.5% SDS, 14 mM 2-mercaptoethanol) plus an equal volume of a phenol and chloroform-isoamyl alcohol mixture (1:1). After

centrifugation, the aqueous phase was recovered and washed with a two-fold volume of the phenol-chloroform-isoamyl alcohol mixture. After centrifugation, a one-tenth volume of 3M sodium acetate, pH 5.2, was added to the aqueous phase to precipitate polysaccharide. The supernatant was collected, and RNAs were precipitated from ethanol as sodium salts. The pellet was resuspended in the TE / HPRI buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 5 units / ml human placenta RNase inhibitor, 1 mM dithiothreitol) and a one-fourth volume of 10 M lithium chloride to precipitate the RNAs as lithium salts. The precipitate was resuspended in TE / HPRI buffer, and the suspension layered on a cushion of a dense solution (5.7 M) of cesium chloride. After centrifugation, the pellet obtained was the stylar total RNA.

Poly(A)+RNA was isolated from the total RNA using a Biomag molecular biology kit for mRNA purification (Perseptive). Double-stranded cDNA was synthesized with a cDNA synthesis kit (Takara).

Isolation and nucleotide sequence analysis of the cDNA clones

The PCR for the S-RNase gene was done with the primers TTTACGCAGCAATATCAG and G(C/T)GGGGGCA(A/G)T(C/T)TATGAA derived from the respective conserved amino acid sequences of *P. pyrifolia* S-RNases, FTQQYQ and FI(D/N)CP(H/R), and Amplitaqgold (Perkin-Elmer) according to the manufacturer's instructions. Each amplified PCR fragment was ligated into pBluescript (Stratagene), then sequenced to confirm that it corresponded to the amino acid sequence of S-RNase. Each fragment was labeled with digoxigenin (Boehringer-Mannheim) and used as a probe for cDNA isolation.

Double-stranded cDNA was inserted into the λZAP II vector (Stratagene) after ligation of the EcoRI / NotI adaptor (Pharmacia) to the cDNAs. Libraries were screened with the digoxigenin-labeled PCR product for each S-RNase gene, and plaques hybridized to these probes were selected. I vectors bearing the

cDNA inserts of S-RNases were converted to plasmids which then were purified by standard procedures. The nucleotide sequence of both strands of double stranded cDNA was determined by the dideoxynucleotide chain-terminating method using Thermosequenase (Amersham) and a Shimadzu DNA sequencer DSQ-1 or Applied Biosystems 373A DNA Sequencer.

Alignment of amino acid sequences and construction of the phylogenetic tree

The amino acid sequences of S-RNases were aligned manually. The amino acid sequence identity among pairs of S-RNases was calculated from the aligned sequences. The neighbor-joining analysis (Saitou and Nei, 1987) of the distance matrix calculated from the amino acid sequence alignment provided the phylogenetic tree. Bootstrap probabilities for clusters were examined. The resampling procedure was repeated 1000 times.

Results

Partial amino acid sequences of five S-RNases from *P. pyrifolia*

To clone the cDNAs of five S- (S_1-, S_3-, S_5-, S_6- and S_7-) RNases, each S-RNase was purified from styles with stigmas by a combination of cation exchange and reversed-phase chromatographies with monitoring the RNase activity. The amino acid sequences of these proteins were then analyzed. [For S_1-, S_6- and S_7-RNases, Mr. Shinkawa in our laboratory purified proteins and analyzed their amino acid sequences (Shinkawa, 1997).] Purified S_5-RNase from styles of the cultivar 'Hosui' (S_3S_5) gave a single peak on reversed-phase chromatography (data not shown) and a single amino acid sequence in a protein

sequencer, which corresponded to that of S_5-RNase (Chapter I; Ishimizu *et al.*, 1996a). The same protein was also purified from styles of the cultivars 'Kosui' (S_4S_5) and 'Okusankichi' (S_5S_7) (data not shown). Moreover, the other four S-RNases were successfully purified. To verify the correspondence of the purified S-RNases to respective *S*-alleles, the *N*-terminal 28, 80, 80, 39, and 18 amino acid residues were sequenced for S_1-, S_3-, S_5-, S_6-, and S_7-RNases, respectively, which were completely identical with the *N*-terminal sequences reported earlier (Chapter I; Ishimizu *et al.*, 1996a).

Each reduced and *S*-carboxymethylated S-RNase was digested with API or V8 or Asn-C. The peptides produced were separated by reversed-phase chromatography (data not shown). Double peaks, probably containing more than one peptide, were collected and rechromatographed. Each peptide was analyzed in a protein sequencer and an amino acid analyzer.

Primary structures of five S-RNases from *P. pyrifolia*

cDNA libraries in λZAPII vectors from styles of the cultivars 'Imamuraaki' (S_1S_6), 'Hosui' (S_3S_5) and 'Okusankichi' (S_5S_7) were constructed and screened with PCR fragments as described in 'Materials and Methods'. Seven to 43 positive plaques were obtained from approximately 20000 plaques by plaque hybridization with each S-RNase probe.

One positive plaque for each S-RNase was picked up and sequenced. The insert of each clone was sequenced. Structural features of the five cloned cDNAs were similar to those of *P. pyrifolia* and *M. domestica* S-RNases cloned so far. Nucleotide sequences of the S_1-, S_3-, S_5-, S_6- and S_7-RNase cDNAs are shown in Figure II-1. All the amino acid residues of S_5-RNase corresponded to those identified by protein or peptide sequencing and by amino acid analysis (Figure II-1(c)). For S_1-, S_3-, S_6- and S_7-RNases, 99, 98, 76, and 61 % of amino acid residues of the respective S-RNases corresponded to those determined by

protein or peptide sequencing (Figure II-1 (a), (b), (d), and (e)). The *N*-terminal amino acid sequences of *P. pyrifolia* S-RNases isolated by two-dimensional gel electrophoresis, reported elsewhere (Chapter I, Ishimizu *et al.*, 1996a), also completely matched the sequences in Figure II-1. The amino acid composition of each S-RNase corresponded to that calculated from the determined amino acid sequence (data not shown). [The accession numbers of DDBJ, EMBL, and GenBank nucleotide sequence databases are AB002139, AB002140, AB002141, AB002142 and AB002143 respectively for the *P. pyrifolia* S_1-, S_3-, S_5-, S_6- and S_7-RNases.]

Comparison of the amino acid sequences of rosaceous S-RNases

Figure II-2 shows the alignment of the amino acid sequences of five S-RNases from *P. pyrifolia* with those of other rosaceous S-RNases. Seventy-six amino acid residues were conserved in rosaceous (*Pyrus* and *Malus*) S-RNases throughout the sequence. Much less conservation was found from the 51st to 66th residue, designated the hypervariable (HV) region. Pairwise amino acid sequence identities (Table II-1) calculated from the aligned sequences were in comparatively narrow ranges of sequence variation (58.8 % to 74.8 %), except between *P. pyrifolia* S_1- and S_4-RNases (90.0 %), and *P. pyrifolia* S_3- and S_5-RNases (95.5%). The respective average amino acid sequence identities among all the pairs of *P. pyrifolia*, *M. domestica* and both alleles were 66.7 %, 67.3 % and 66.1 %. Some interspecific sequence similarities of the S-RNases were higher than those within a species (Table II-1). The phylogenetic tree (neighbor-joining tree) of the amino acid sequences of 12 maloideous S-RNases also indicates the discrepancy between the S-RNase sequence similarity and the taxonomic relationship in *Pyrus* and *Malus* (Figure II-3), suggesting that the S-RNase polymorphism predates the divergence of *Pyrus* and *Malus*.

The motifs encompassing two histidine residues essential for RNase activity (Kawata *et al.*, 1990) (DKLFTVHGLWPS, WxKHGxC) were conserved in rosaceous S-RNases. Eight cysteine residues that form four disulfide bonds (Chapter III, Ishimizu *et al.*, 1996b) also are completely conserved. The number of potential glycosylation sites with the consensus sequence Asn-Xaa-Ser / Thr (Xaa is not Pro and Asp) are two (*P. pyrifolia* S_3- and S_5-RNases) to six (*M. domestica* S_7-and S_9-RNases). These sites are dispersed throughout the sequence, including the HV region. Of them, Asn 121 in the sequence KQNVS is the only conserved site.

The high amino acid sequence similarities between *P. pyrifolia* S_1-and S_4-RNases (90.0 %) and between *P. pyrifolia* S_3-and S_5-RNases (95.5 %) are remarkable (Figure II-2 and Table II-1). Positions of the amino acid substitutions in these pairs are shown in Figure II-4. Sites of the substitutions between *P. pyrifolia* S_1- and S_4-RNases are spread over the entire region, but two consecutive amino acid substitutions are present in the HV region. In *P. pyrifolia* S_3- and S_5-RNases, they are restricted to the region consisting of 70 amino acid residues (amino acid numbers 21-90), which includes the HV region.

The flanking region of the coding region of rosaceous S-RNases

The sequence around the first ATG codon (ATTCAATG) in the 5' end was well conserved in the cDNAs of the rosaceous S-RNases, except for *P. pyrifolia* S_3-RNase in which the substitution was ATTCACTA. The second ATG codon was conserved in all rosaceous S-RNases. Which ATG codon coded the initiator methionine could not be determined. The expression of *P. pyrifolia* S_3-RNase did not differ from the expressions of the other *P. pyrifolia* S-RNases, as seen from the S-RNase spots from results of two-dimensional gel electrophoresis of stylar protein (Chapter I; Ishimizu *et al.*, 1996a) and the yield of purified S-RNases from stylar extracts (unpublished data). The putative signal peptide is shown in italics

in Figure II-4. These are typical signal peptide sequences composed of a central hydrophobic core flanked by polar residues (Watson, 1984). The signal sequence cleavage site conforms to the (-3, -1)-rule (von Heijine, 1986).

P. pyrifolia S_1-RNase had a dinucleotide repeat sequence (CA, AT) in the 3' flanking region. *P. pyrifolia* S_4-RNase (Norioka *et al.*, 1996), *M. domestica* S_5-, S_9-, (Janssens *et al.*, 1995), and S_c- and S_f-RNases (Sassa *et al.*, 1996) also had dinucleotide repeat sequences in equivalent positions. Relatively short stretches of dinucleotides, apparently remnants of these long repeated sequences, are present in *P. pyrifolia* S_2- (Norioka *et al.*, 1996), S_3-, S_5- and S_7-RNases and *M. domestica* S_7-RNase (Janssens *et al.*, 1995).

Discussion

S-RNase is associated with GSI as a stylar component encoded by the *S*-locus. I determined the primary structures of five S-RNases from *P. pyrifolia* by amino acid sequencing of the purified proteins and nucleotide sequencing of their cDNAs. The amino acid sequences deduced from nucleotide sequences completely matched those determined by protein sequencing in this and a previous chapter (Ishimizu *et al.*, 1996a). Then the primary structures of rosaceous S-RNases were characterized based on their primary structures.

The *N*-terminal amino acid sequences of S-RNases from *Pyrus ussuriensis* (Ussurian pear), *Pyrus communis* (European pear), and *Prunus dulcis* (almond) have been reported (Tomimoto *et al.*, 1996; Tao *et al.*, 1997) (Figure II-2). The *N*-terminal amino acid sequences of S-RNases of *P. ussuriensis* and *P. communis* are similar to those of *P. pyrifolia* and *M. domestica*, whereas those of *P. dulcis* are distinct. The Rosaceae generally is divided into four subfamilies: the

Maloideae (=Pomoidaeae) (x=17), Prunoideae (=Amygdaloideae) (x=8), Rosoideae (x=7), and Spiraeoideae (x=9). The Maloideae comprises *Pyrus*, *Malus*, and *Cydonia*, and the Prunoideae the genus *Prunus*. Chemotaxonomic data on the presence of flavone C-glycosides and phylogenetic analysis data on the *rbc*L gene sequences are consistent with these divisions, and indicate that Maloideae arose by autopolyploidy or allopolyploidy from primitive forms of the Spiraeoideae (or Prunoideae) (Challice, 1974; Morgan, *et al.*, 1994). Species within the Maloideae are closely related, intergeneric crosses and intergeneric grafts between *Pyrus* and *Cydonia*, and between *Pyrus* and *Malus* having been reported (Shimura *et al.*, 1983; Westwood *et al.*, 1989; Banno *et al.*, 1993). These taxonomic facts are consistent with the division that the amino acid sequences of S-RNases of Maloideae are more closely related to each other than to those of the Prunoideae (Figure II-2 and Table II-1). The phylogenetic tree suggests that *P. pyrifolia* and *M. domestica* S-RNases are closely related and that S-RNase polymorphism predates the divergence of *Pyrus* and *Malus*. Sassa *et al.* (1996) also reported this observation from 6 primary structures of maloideous S-RNases. A similar trans-species evolution was reported in the solanaceous S-RNases (Ioerger *et al.*, 1990) and in the primate major histocompatibility complex (MHC) proteins which was pointed out that balancing selection operated (Figueroa *et al.*, 1988; Lawlor *et al.*, 1988).

Completely conserved amino acids make up ~38 % and ~17 % respectively of the total residues of the 12 rosaceous S-RNases (Figure II-2) and 19 solanaceous S-RNases (Tsai *et al.*, 1992). The lowest score for pairwise amino acid sequence identity in rosaceous S-RNases is 58.8 % (*P. pyrifolia* S_1- and S_2-RNase and *P. pyrifolia* S_2- and S_4-RNase) (Table II-1), whereas that for the solanaceous S-RNases is 38.7 % (Tsai *et al.*, 1992). The similarities range from 60 % to 70 % for most rosaceous S-RNase pairs (Table II-1), whereas they range from 40 % to 50 % for most solanaceous S-RNase pairs (Tsai *et al.*, 1992).

Phylogenetic tree of the RNase T_2 family enzymes indicates that the rosaceous S-RNases are obviously distinct from the solanaceous S-RNases and suggests that these two S-RNase groups evolved independently (Xue *et al.*, 1996; unpublished data), but it is not clear whether the origin of these S-RNases is the same. This problem was discussed well by Xue *et al.* (1996).

The sequence identity between *P. pyrifolia* S_3- and S_5-RNases is very high (95.5%), and amino acid substitutions are found only in the stretch of 70 amino acids including the HV region (amino acid numbers 21-90) (Figure II-4). Such a pair has not been found before. McCubbin *et al.* (1997) reported experimental evidence that S-RNase is a recognition molecule that interacts with its counterpart. Amino acid substitution in this restricted region between *P. pyrifolia* S_3- and S_5-RNases appears to be sufficient to discriminate between S_3- and S_5-pollen and to trigger the self-incompatible reaction.

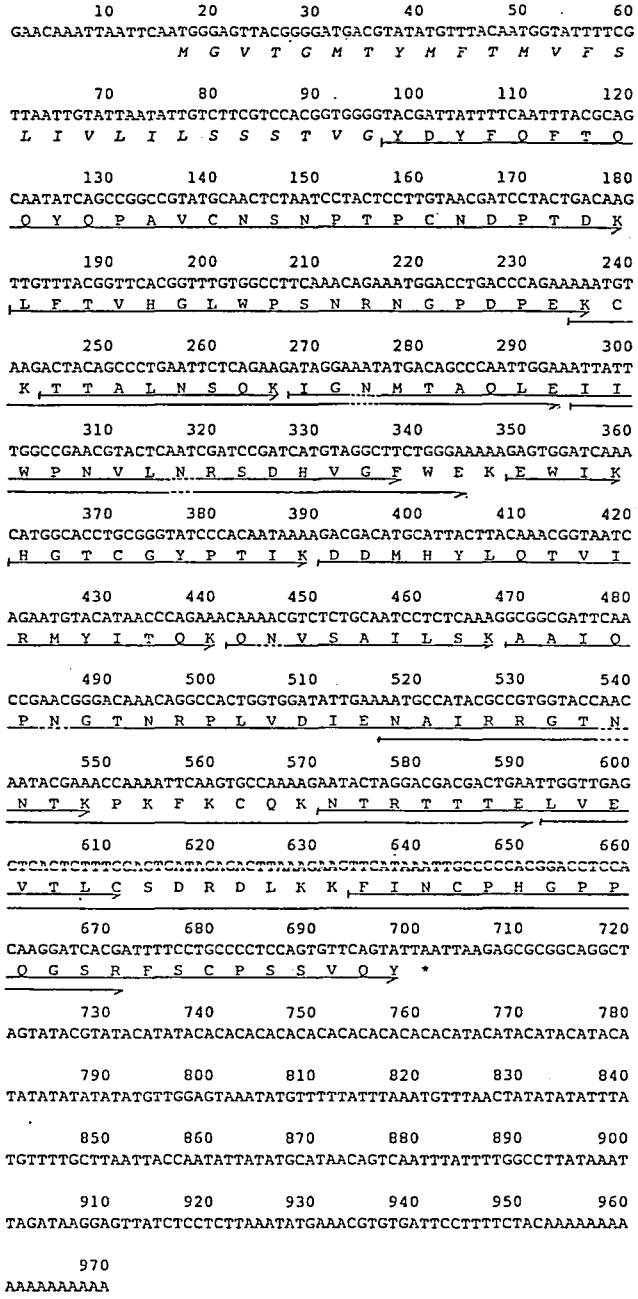
The pairs of *P. pyrifolia* S_1- and S_4-RNases (Figure II-4) and *Solanum chacoense* S_{11}- and S_{13}-RNases (Saba-El-Leil *et al.*, 1994) also showed high sequence homology. Amino acid substitution sites in these pairs are scattered throughout the sequences unlike in *P. pyrifolia* S_3- and S_5-RNases and only a few amino acid substitutions are present in the HV region. Recently, it has been reported that four amino acid substitutions in the HV region between *Solanum chacoense* S_{11}- and S_{13}-RNases are necessary and sufficient to discriminate between S_{11}- and S_{13}-pollen (Matton *et al.*, 1997). This result suggests that only two of the 19 substitutions between *P. pyrifolia* S_1- and S_4-RNases, which are within the HV region, may be responsible for the *S*-allele-specific recognition (Figure II-4). But, other experiments of domain exchange using transgenic plants in solanaceous S-RNases suggested that *S*-allele recognition site(s) is not restricted to the HV region (Kao and McCubbin, 1996; Zurek *et al.*, 1997), which conflicts with the result of *Solanum chacoense* S-RNases. Accordingly, it has to

be carefully examined where the *S*-allele recognition site(s) of *P. pyrifolia* S-RNases is.

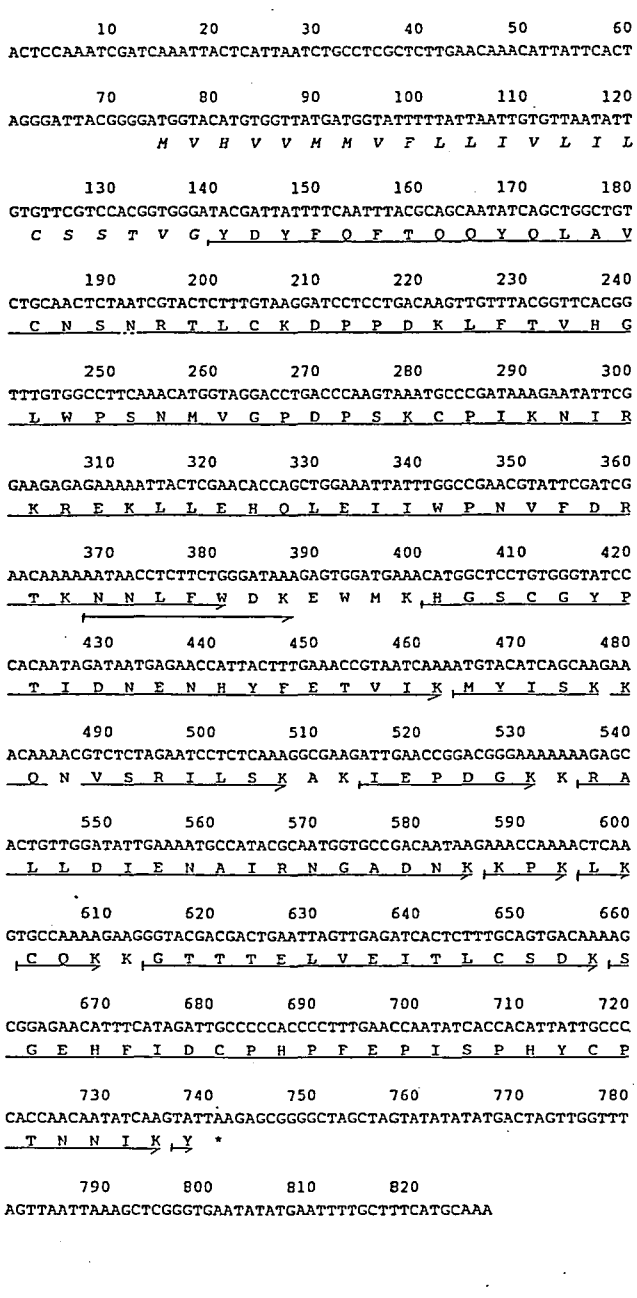
I cloned the cDNA of *P. pyrifolia* S_5-RNase from the cultivar 'Hosui' (S_3S_5). Recently the cDNA clone of *P. pyrifolia* S_5-RNase was isolated from the cultivar 'Kosui'(S_4S_5) (Sassa and Hirano, 1997). Although the genetic relationship between the two cultivars is not clear because of lack of knowledge on the genetic background of 'Hosui' (Machida *et al.*, 1982), these sequences are identical at the amino acid and nucleotide level. Each SLG^{24} of the self-incompatible *Brassica campestris* from Japanese and Turkish populations also has the same nucleotide sequence (Matsushita *et al.*, 1996). In this respect, *M. domestica* S_9-RNase (Janssens *et al.*, 1995) from 'Queen's Cox', mainly cultivated in Western nations, and *M. domestica* S_c-RNase (Sassa *et al.*, 1996) from 'Fuji', mainly cultivated in Japan, are an interesting pair. The identity of the amino acid sequences of the two S-RNases is 100 % and that of the nucleotide sequence 98.6 %. Differences in the nucleotide sequence were found only in the 3'-flanking region. The main difference is in the numbers of GT and AT repeat sequences. A constraint on nucleotide substitution may operate on the coding region of S-RNase to maintain allelic identity. A similar observation was reported for genes encoding the SI-associated proteins of *Papaver rhoeas* (Walker *et al.*, 1996). These dinucleotide repeat sequences (GA, or CA, or AT) also are present in almost all rosaceous S-RNases cloned so far. The GT or CA sequence repeat forms left-handed Z-DNA, which is considered to function in transcriptional control, recombination, and the organization of the nuclear chromatin structure (Rich *et al.*, 1984). The role of the dinucleotide repeat sequence in the S-RNase gene, however, has yet to be investigated.

Figure II-1. **Nucleotide and amino acid sequences of five S-RNases from *P. pyrifolia*.** The deduced amino acid sequence of each S-RNase is shown under the nucleotide sequence of its cDNA. Sequences determined by protein or peptide sequence analyses are underlined with arrows. A dashed line denotes an unidentified PTH-amino acid. Putative leader peptides are shown in Italics. The translation termination codon is shown by an asterisk. (a) S_1-RNase. (b) S_3-RNase. (c) S_5-RNase. (d) S_6-RNase. (e) S_7-RNase. The accession numbers of DDBJ, EMBL, and GenBank nucleotide sequence databases are AB002139, AB002140, AB002141, AB002142 and AB002143 respectively for the *P. pyrifolia* S_1-, S_3-, S_5-, S_6- and S_7-RNases.

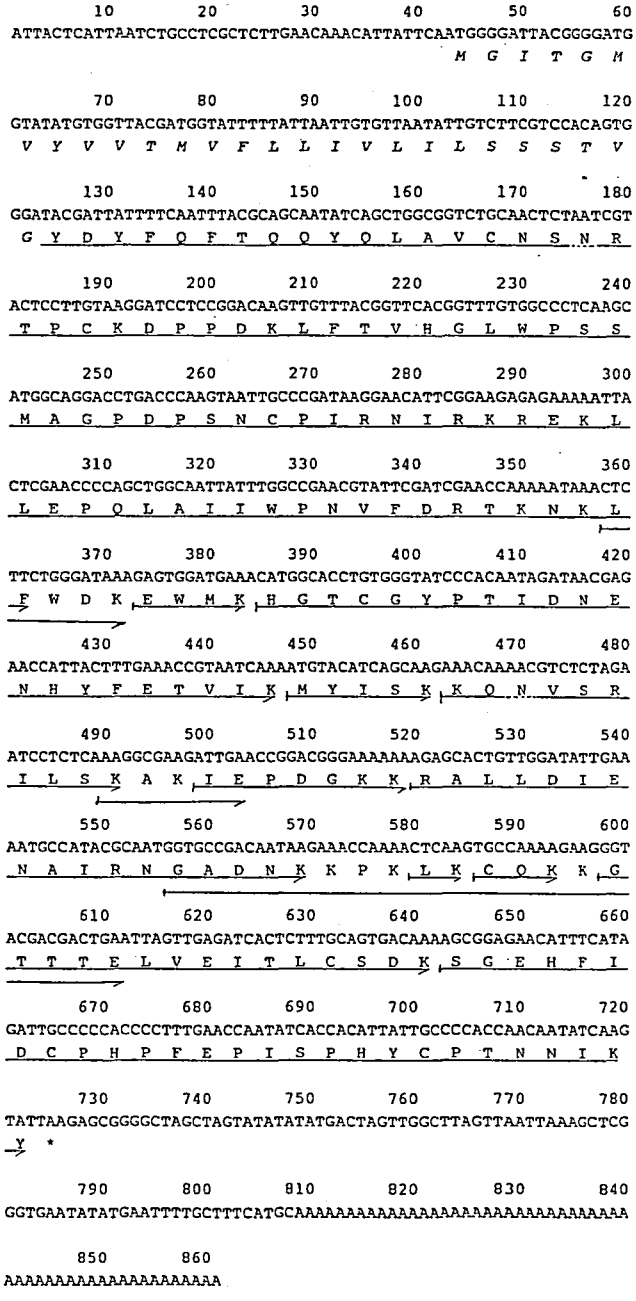
(a)



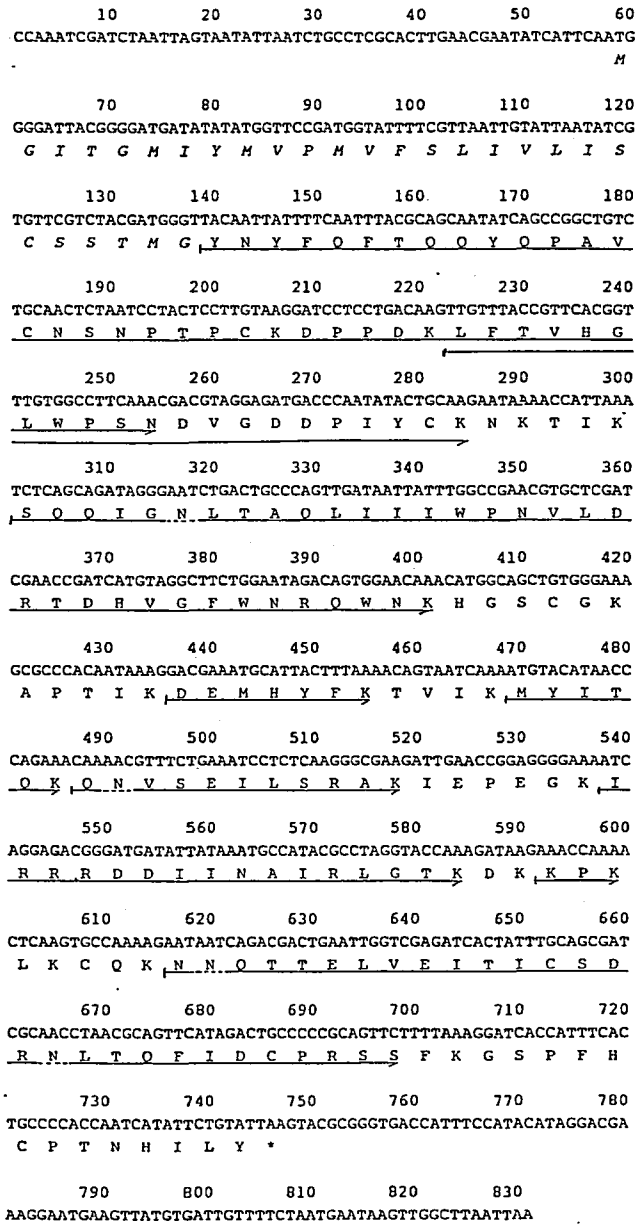
(b)



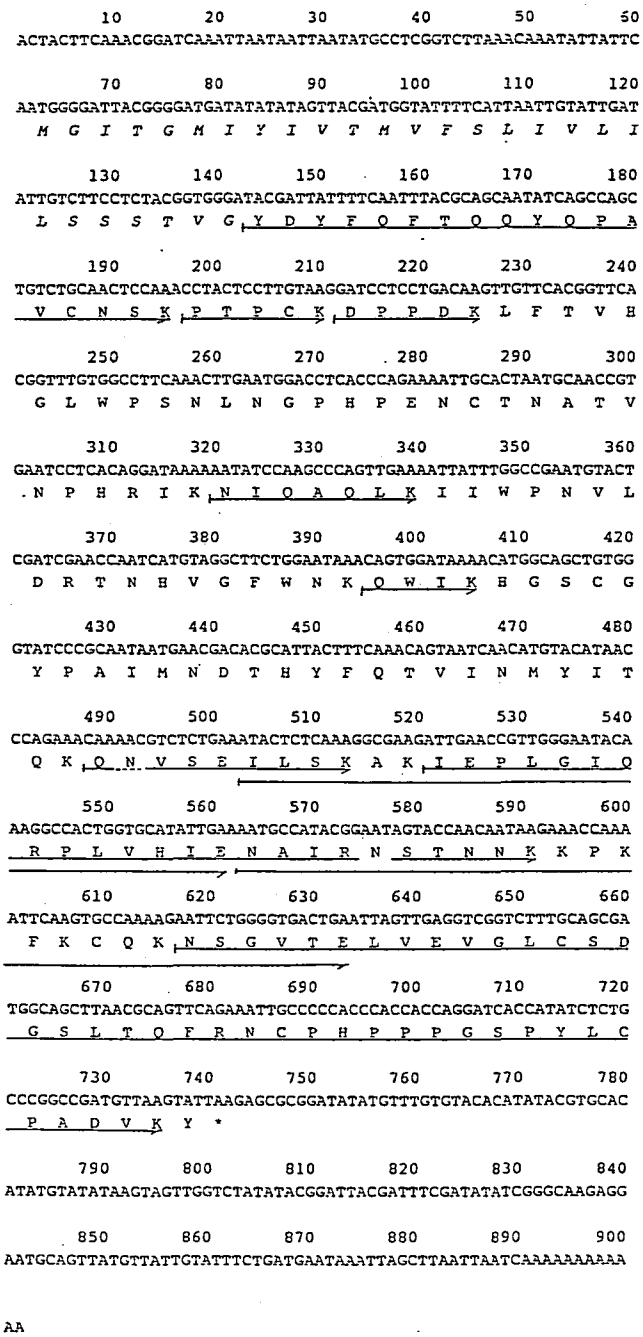
(c)



(d)



(e)



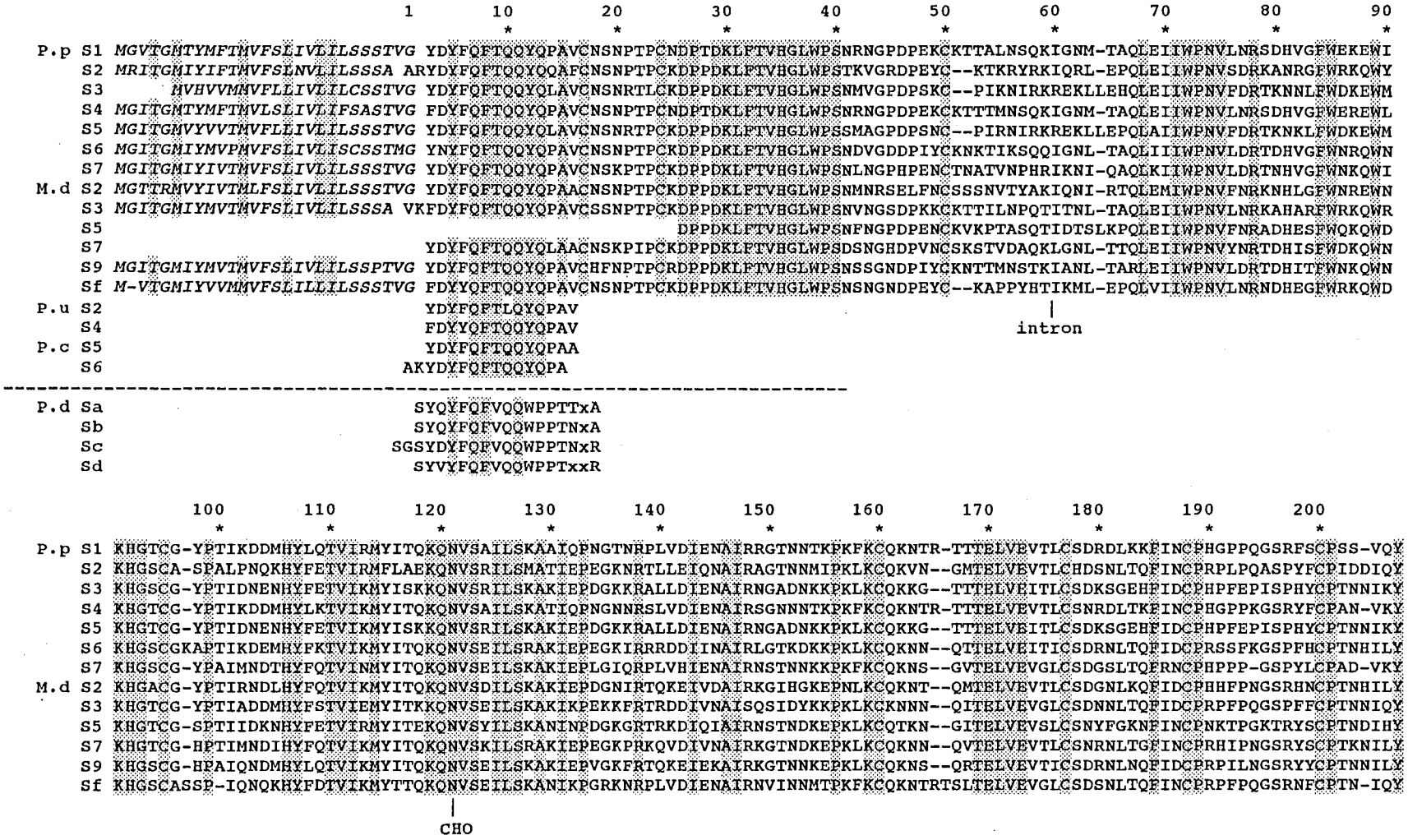


Figure II-2. **Alignment of the amino acid sequences of 21 rosaceous S-RNases.** Sequences were aligned manually. Leader peptides of S-RNases are shown in Italics. Gaps are marked by dashes. The conserved amino acid residues in rosaceous S-RNases are shaded. The only conserved *N*-glycosilation site in the rosaceous S-RNases is marked with CHO. The site of the intron is shown under the sequences. Numbering starts at the *N*-terminus of *P. pyrifolia* S_2-RNase and *M. domestica* S_3-RNase. Abbreviations : P.p S_1, S_2, S_3, S_4, S_5, S_6, and S_3, S_1-, S_2-, S_3-, S_4-, S_5-, S_6-, and S_7-RNases from *P. pyrifolia* (Norioka *et al.*, 1995; this chapter); M.d S_2, S_3, S_5, S_7, S_9, and S_f, S_2-, S_3-, S_5-, S_7-, S_9-, and S_f-RNases from *M. domestica* (Broothaerts *et al.*, 1995; Janssens et al., 1995; Sassa *et al.*, 1996); P.u S_2 and S_4; S_2- and S_4-RNases from *P. ussuriensis*. (Torimoto *et al.*, 1997); P.c S_5 and S_6, S_5- and S_6-RNases from *P. communis* (Torimoto *et al.*, 1997); and P.d S_a, S_b,S_c, and S_d, S_a-, S_b-, S_c-, and S_d-RNases from *Prunus dulcis* (Tao *et al.*, 1997).

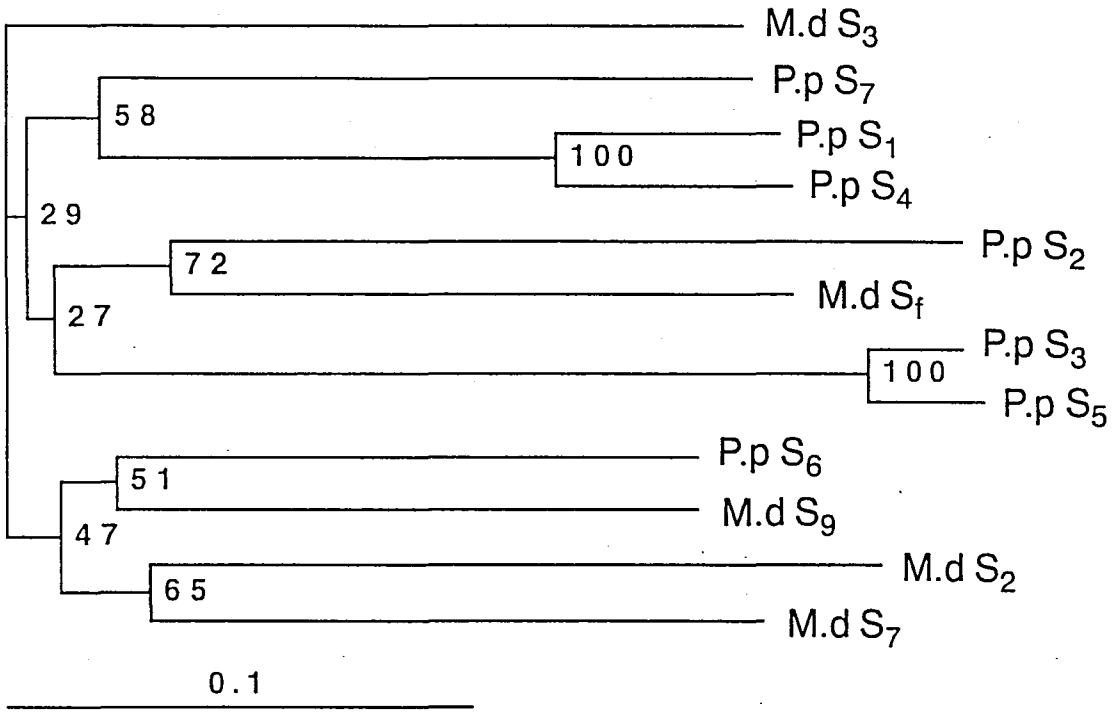
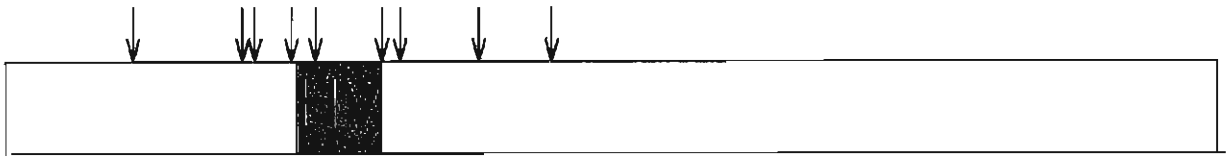


Figure II-3. **Neighbor-joining phylogenetic tree of maloideous S-RNases.** Bootstrap probabilities for clusters are shown as percentages. The bar under the tree represents the number of nucleotide substitutions.

P. pyrifolia S_3- and S_5-RNases (95.5%)



P. pyrifolia S_1- and S_4-RNases (90.0%)

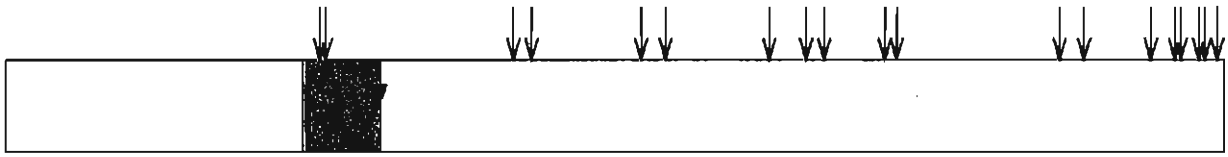


Figure II-4. **Amino acid substitution sites in highly homologous pairs of S-RNases.** The HV regions in rosaceous S-RNases are boxed. Amino acid substitution sites in each pair are marked with arrows.

Table II-1. **Pairwise amino acid sequence identities of *P. pyrifolia* and *M. domestica* S-RNases listed in Figure II-4, with *M. domestica* S_5-RNase omitted.** Values show the percent of amino acid identity.

		P. pyrifolia							*M. domestica*				
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_2	S_3	S_7	S_9	S_f
P. pyrifolia	S_1		58.8	62.1	90.0	61.1	67.0	71.1	63.9	68.1	66.8	67.8	67.8
	S_2			62.4	58.8	61.9	62.3	64.0	59.1	61.1	62.1	62.6	68.0
	S_3				62.1	95.5	64.5	63.4	62.4	62.7	61.9	63.9	60.9
	S_4					61.6	69.0	70.4	63.9	66.7	68.3	68.3	67.8
	S_5						63.5	63.4	62.4	62.3	62.9	63.4	60.4
	S_6							66.8	66.3	70.6	71.3	74.8	64.2
	S_7								65.2	65.0	67.2	69.7	68.8
M. domestica	S_2									65.0	71.6	69.7	62.1
	S_3										66.0	68.0	67.3
	S_7											73.6	64.0
	S_9												66.0
	S_f												

Chapter III
Location of disulfide bonds in S-RNases

Introduction

In the previous chapters, I identified seven S-RNases associated with GSI of *P. pyrifolia* and determined their primary structures. In this and next chapters, I analyzed post-translational modification of S-RNases.

About 50 S-RNase genes have been cloned so far from rosaceous and solanaceous plants and their nucleotide sequences determined. Sequence alignment of their putative amino acid sequences show that they have about 20 completely conserved amino acid residues, including two catalytic histidine (Kawata *et al.*, 1990) and eight half-cystine residues (loerger *et al.*, 1991; Chapter II, Ishimizu *et al.*, 1998b). *N. alata* S_2-, S_3-, and S_6-RNases had one, two, and two free cysteine residues, respectively, and all three S-RNases had four disulfide bonds (Ishimizu *et al.*, 1995). Cys 95 in *N. alata* S_6-RNase was identified as a free cysteine residue (Ishimizu *et al.*, 1995). But it is not clear that location of other free cysteine and disulfide bonds. These half-cystines have been found in some plant RNase T_2-type enzymes unrelated to GSI, such as the ribonuclease encoded by RNS genes 1, 2 and 3 from *Arabidopsis* (Taylor *et al.*, 1993; Bariola *et al.*, 1994), RNase LE from cultured tomato cell (Jost *et al.*, 1991) and non-S-RNase from *P. pyrifolia* (Norioka *et al.*, 1996). In contrast, the fungus RNase T_2-type enzymes RNase T_2 (Kawata *et al.*, 1989) and RNase Rh (Horiuchi *et al.*, 1988) have ten half-cystine residues, only four of which are shared with the higher plant enzymes. These facts suggest that the disulfide bond location differs

considerably between higher plant and fungus enzymes, but this has yet to be fully investigated for RNase T_2-type enzymes.

Cysteine thiols and cystine disulfides often are very important factors affecting protein structure and function. Chemical modification of the purified protein is needed to clarify the state of the half-cystine residues because no information on this event can be obtained from the nucleotide sequence of the gene. When thiol groups and disulfide bonds coexist in a protein such as *N. alata* S_6-RNase, the thiol-disulfide exchange reaction is accelerated in alkaline or strongly acidic media. In fact, *N. alata* S_2-RNase, which had four disulfide bonds and one free cysteine (Ishimizu *et al.*, 1995), suffered from the thiol-disulfide exchange reaction in alkaline media (Oxley and Bacic, 1995). This means that experiments for determining disulfide bonds in the presence of free thiol groups must be carefully designed to suppress the exchange reaction. I devised a new method to analyze the cysteine and cystine residues in a protein. Briefly, thiol groups first are pyridylethylated by 4-VP (4-vinylpyridine) at an acidic pH to suppress the thiol-disulfide exchange, resulting in the production of PE-protein. The PE-protein is analyzed for determining the location of the free cysteines and disulfide linkages in a protein.

I used this method to analyze the cysteine and cystine residues in *N. alata* S_6-RNase and *P. pyrifolia* S_4-RNase. S_6-RNase and S_4-RNase have eight half-cystine residues conserved in all S-RNases (Figure III-1) (Anderson *et al.*, 1989; Norioka *et al.*, 1996). S_6-RNase has two additional half-cysteine residues (Cys77 and Cys95), Cys95 is free cysteine (Ishimizu *et al.*, 1995). No investigation has been made of the state of the half-cystines of S_4-RNase. I here describe the identification by our new method of the free cysteine residues and disulfide linkages in the two S-RNases and discuss the common disulfide bridge motif in the S-RNases. I also report the sugar compositions of the two S-RNases as deduced from the glycopeptide molecular masses in their API-digests.

(Of the context of this chapter, *N. alata* S_6-RNase was purified by Drs. Lush, M, Anderson, M.A., Opat, A and Prof. Clarke, A.E. of University of Melbourne, and purification of S_4-RNase and fragmentation of PE-S_4-RNase were carried out by Dr. Norioka, S. of our laboratory.)

Materials and Methods

Materials

N. alata S_6-RNase was purified from styles of the S_6S_6 homozygote of *N. alata*, as reported elsewhere (Jahnen *et al.*, 1989). *P. pyrifolia* S_4-RNase was obtained by a series of CM-cellulose and reversed-phase chromatographies (Chapter I, Ishimizu *et al.*, 1996a). The reagents were ovalbumin (Sigma Chemical), 4-VP and API (Wako Pure Chemicals), 2-mercaptoethanol (Katayama Chemicals). All other chemicals were of the highest grade available commercially.

S-Pyridylethylation of thiol groups at an acidic pH

The protein was dissolved in 60 μl of 0.1M sodium acetate (pH 4.5 to 5.5) or 0.1 M Mes / NaOH buffer (pH 6.0 to 6.5) containing 6 M guanidine hydrochloride, then a specified amount of 4-VP diluted with acetonitrile was added to the solution. After replacing the air by flushing the tube with N_2 gas, the tube was kept at an ambient temperature in the dark for 4 hr. PE-protein was desalted in a Fast Desalting column (3.2 x 100 mm) equilibrated with 0.05% trifluoroacetic acid or 10 mM sodium phosphate buffer (pH 6.5) using the Pharmacia SMART system.

Reduction and *S*-carboxymethylation

Reduction and *S*-carboxymethylation were done by the method of Crestfield *et al.* (1963). The reduced, *S*-carboxymethylated protein (RCM-protein) was desalted as described above.

Digestion of PE-protein with API

PE-protein was digested with API in 10 mM sodium phosphate buffer, pH 6.5, at 37°C for 1 hr at the enzyme / substrate ratio of 1 / 50 (mol / mol).

LC/ESI-MS of the API-digest

ESI-MS was done in a Finnigan MAT TSQ 7000 mass spectrometer equipped with a Finnigan MAT atmospheric pressure ionization interface operating in the electrospray ionization mode with 4.5 kV of needle voltage and a heated capillary temperature of 200 °C. A Michrom BioResources Ultrafast Microprotein Analyzer (Pleasanton) fitted with a 10nm PLRP-S column (1.0 x 150 mm) was coupled directly to the mass spectrometer via an interface. The API-digest was applied to the column equilibrated with 0.1 % formic acid in 98% water and 2% acetonitrile. The organic solvent (0.095% formic acid in 2% water and 98% acetonitrile) gradient was developed linearly to 50% for 50 min at a flow rate of 50 ml / min. Each peptide eluted was injected to the mass spectrometer via the interface. The quadrupole was scanned over 330-2500 Da every 3 seconds. The series of m / z values of the multiple charged peptides in each peak were deconvoluted to a given molecular mass.

Amino acid and sequence analyses

The proteins or peptides were hydrolyzed in evacuated tubes in twice-distilled 5.7 M HCl containing 0.2% phenol for 24 hr at 110 °C. Amino acid analysis of the hydrolysate was done with a Hitachi L-8500S amino acid analyzer. Sequence

analyses of the proteins or peptides were done with a gas-phase protein sequencer (Applied Biosystems model 470A equipped with an on-line 120A Pth (phenylthiohydantoin) analyzer). Pth-Xaa were analyzed with a modified isocratic elution system (Tsunasawa *et al.*, 1985).

Results

Optimization of pyridylethylation with 4-VP under an acidic condition

Chicken ovalbumin has four free cysteine residues and one disulfide bond (Thompson and Fisher, 1978). To optimize thiol-specific pyridylethylation at an acidic pH, 500 pmol samples of ovalbumin were allowed to react with various amounts of 4-VP for 4 hr at pH 4.5-6.5. The reaction was stopped by desaltation, then, after reduction and *S*-carboxymethylation, the number of PE-Cys in the protein was estimated by amino acid analysis. When a 50- to 100-fold molar excess of 4-VP was added to the sample solution at pH 6.0 or 6.5, about 95% of the four cysteine residues in the ovalbumin was S-pyridylethylated (Figure III-2). S-pyridylethylation was not complete even when a large excess of 4-VP and a prolonged reaction time were used (data not shown). As shown in Table III-1, no serious side reaction occurred at any pH for any other residues during the pyridylethylation process. Consequently, the conditions considered optimal for thiol-specific pyridylethylation were incubation of the protein with a 50-fold molar excess of 4-VP in 0.1M Mes / NaOH buffer, pH6.0, for 4 hr at room temperature.

Number of free cysteines in the S-RNases

When 500 pmol of *N. alata* S_6-RNase was incubated with 4-VP under the conditions described above, 1.9 mol of PE-Cys and 7.9 mol of CM-Cys per 1 mol

of protein were detected by amino acid analysis after reduction and S-carboxymethylation (Table III-2), indicative that S_6-RNase has two free cysteine and four cystine residues. Similar results had been obtained by 5,5'-dithio-bis(2-nitrobenzoic acid) titration (Ishimizu *et al.*, 1995). In contrast, amino acid analysis of the reduced and *S*-carboxymethylated PE-S_4-RNase (RCM-PE-S_4-RNase) of *P. pyrifolia* showed that this protein has four cystine residues and no free thiol group (Table III-2).

Location of free cysteine residues and disulfide bonds in *N. alata* S_6-RNase

The location of the cysteine and cystine residues in the protein was determined at an acidic pH. A 300 pmol portion of PE-S_6-RNase was digested with API at pH 6.5, and the digest subjected to LC/ESI-MS, resulting in its separation into major thirteen fragments (Figure III-3). As shown in Table III-3, the molecular mass of the peak eluted at 37.5 min (2613.7 Da) was consistent with the calculated mass of K1-peptide (2613.1 Da), including the disulfide bond between Cys16 and Cys21 (with 0.023% error), the K-peptides being numbered in order from the *N*-terminus of the protein. The peak eluted at 33.1 min gave several deconvoluted mass values, because of the heterogeneity of the sugar moiety (Figure III-4). *N*-terminal amino acid sequence analysis revealed that the peak consisted of K2- and K7-peptides, and Cys94 was identified as a dithiothreitol adduct of Pth-dehydroalanine in the fourth cycle of Edman degradation (Table III-4). This product was thought to be derived from a cysteine residue produced by the reductive cleavage of a disulfide bond in the reaction cartridge of the sequencer. Pth-cysteine and a trace amount of Pth-cystine also were detected in this cycle. In fact, when cystine was placed in the sequencer, these three Pth-Xaa were detected (data not shown). In contrast, in the fifth cycle Cys95 was identified by the sequencer as Pth-PE-cysteine. These results

indicate that Cys94 was linked to Cys45 by a disulfide bond and that Cys95 existed as a free cysteine residue. The observed mass of the peak eluted at 3.6 min (671.5 Da) agreed with the calculated mass of the K5-peptide pyridylethylated at Cys77 (671.8 Da) with 0.045% error (Table III-3). Similarly, the respective peaks eluted at 13.9 min (1330.2 Da) and 28.3 min (2780.6 Da) were identified as K13-K16-peptide (1330.5 Da) and K14-K15-peptide (2780.1 Da), in which there were two disulfide bonds at Cys153-Cys182 and Cys165-Cys176 (Table III-3). Other peaks containing no cysteine residue also were assigned to appropriate K-peptides on the basis of their molecular masses, except for the peak eluted at 4.9 min which was assigned to the K11-peptide bearing an *N*-linked glycan at Asn138 by *N*-terminal sequence analysis (data not shown). Only the K12-peptide was not detected on the chromatogram (Figure III-3), probably because it was eluted at void volume. Because the average deviation of the observed masses was 0.029%, the assignment of the individual peptides was sufficiently reliable. Moreover, thiol-disulfide exchange seemed to be suppressed throughout all the processes because no disulfide bond or PE-cysteine, other than those shown in Table III-3, were found. These results indicate that *N. alata* S_6-RNase has two free cysteines, Cys77 and Cys95, and four disulfide bonds (Cys16-Cys21, Cys45-Cys94, Cys153-Cys182 and Cys165-Cys176).

Glycosylation in *N. alata* S_6-RNase

S_6-RNase had four potential *N*-glycosylation sites, Asn27, Asn37, Asn138 and Asn150 (Anderson *et al.*, 1989). Two of the four sites, Asn27 and Asn37, exist in the K2-K7-peptide, and the other two, Asn138 and Asn150, respectively are present in the K11- and K13-K16-peptides (Table III-3). As shown in Figure III-4 and Table III-5, six major deconvoluted peaks were observed for the K2-K7-peptide and two major ones for the K11-peptide. The K13-K16-peptide showed

only a single peak and Pth-asparagine was clearly detected at the fourth cycle as analyzed with the protein sequencer, indicative that it had no *N*-glycan at Asn150. For the K11-peptide, when one of the two deconvoluted peaks (1929.8 Da) was tentatively assigned as a glycopeptide bearing three *N*-acetylhexosamine, three hexose and one pentose residue ($HexNAc_3Hex_3Pen_1$) (Table III-5), the calculated mass of the peptide (1929.9 Da) was consistent with the observed mass, with 0.005% error. Similarly, the glycan of the 2091.7 Da peak is presumed to be the same as that of the 1929.8 Da peak, but with an additional hexose residue ($HexNAc_3Hex_4Pen_1$) (Table III-5). Based on the structures of the *N*-linked glycans for the tobacco S_1- and S_2-RNases (Oxley and Bacic, 1995) and the putative structure for S_6-RNase (Ishimizu *et al.*, 1995), Asn138 is presumed to have the hybrid-type glycan composed of the sugar core (two *N*-acetylglucosamine and three mannnose residues), as well as an *N*-acetylglucosamine and a xylose residue (plus a mannose residue).

Assuming that the highest peak (7778.2 Da) for the K2-K7 peptide (Figure III-4) was the glycopeptide-bearing $HexNAc_6Hex_6Pen_1$, the calculated mass (7775.8 Da) was very consistent with the observed one. The sugar compositions of the other peaks recorded for the K2-K7-peptide likewise were deduced, as shown in Table III-5. In the 7778.2 Da peak, both the glycosylation sites Asn27 and Asn37 may have the *N*-linked glycan composed of $HexNAc_3Hex_3$, and one of the two sites has an additional pentose residue, resulting in the sugar composition of $HexNAc_6Hex_6Pen_1$. The 7912.0, 7980.8 and 8115.8 Da peaks corresponded respectively to the 7775.8 Da peptide plus an additional pentose, an *N*-acetylhexosamine, or both. In the 7722.2 Da peak, if each site has the sugar core $HexNAc_2Hex_3$, six sets of glycans are possible at the two sites: $HexNAc_3Hex_3$ and $HexNAc_3Hex_6$, $HexNAc_3Hex_4$ and $HexNAc_3Hex_5$, $HexNAc_2Hex_3$ and $HexNAc_4Hex_6$, $HexNAc_2Hex_4$ and $HexNAc_4Hex_5$, $HexNAc_4Hex_3$ and $HexNAc_2Hex_6$, $HexNAc_4Hex_4$ and $HexNAc_2Hex_5$. The 8059.6

Da peak corresponded to the 7722.2 Da peptide plus an additional *N*-acetylhexosamine and pentose residue.

Disulfide bonds in *P. pyrifolia* S_4-RNase

LC/ESI-MS analysis of the API-digest of *P. pyrifolia* PE-S_4-RNase (200 pmol) was done as described for the S_6-RNase (Figure III-5) (Fragmented PE-S_4-RNase was prepared by Dr. Norioka, S. of our laboratory.). Each peak obtained was assigned to an appropriate K-peptide (K-peptides being numbered in order from the *N*-terminus of the S_4-RNase) on the basis of its observed molecular mass (Table III-6). Peaks with a disulfide linkage were obtained at retention times of 29.5 min (K1), 13.3 min (K3-K6), 14.0 min (K14-K17), 17.3 min (K14-K17+K18) and 23.0 min (K15-K16). The peak eluted at 17.3 min was the K14-K17-peptide linked to the K18-peptide. All eight cysteine residues in the four peptides K1, K3-K6, K14-K17 and K15-K16 formed a disulfide bridge with their counterparts and were not pyridylethylated, which is consistent with results of the amino acid analysis of the RCM-PE-S_4-RNase. Because each of the peaks eluted at 20.9, 36.8 and 37.1 min showed several deconvoluted masses (Table III-7), they seemed to be glycopeptides. *N*-terminal sequence analysis showed that the peak eluted at 20.9 min was the K10-peptide and the peaks eluted at 36.8 and 37.1 min were derived from the K5-peptide (denoted the K5a- and K5b-peptides) (Table III-7). The two peaks eluted at 21.6 and 22.1 min did not correspond to any K-peptide in terms of molecular mass. By means of *N*-terminal sequence analysis both peaks were assigned to glycopeptides derived from the K11-peptide (designated the K11a- and K11b-peptides) although they did not show a series of deconvoluted peaks (Table III-7). No peaks corresponding to the K12-, K13- and K18-peptides were recorded on the chromatogram because their masses were below the minimum limit of the scan range (<330 Da). The other peaks were assigned to the appropriate K-peptides. The average deviation

was 0.023%, and no disulfide bond or PE-cysteine other than the ones listed in Table III-6 were detected. In conclusion, there are four disulfide linkages (Cys15-Cys22, Cys48-Cys92, Cys156-Cys195 and Cys172-Cys183) in *P. pyrifolia* S_4-RNase.

Glycosylation in *P. pyrifolia* S_4-RNase

P. pyrifolia S_4-RNase has five potential *N*-glycosylation sites (Asn60, Asn74, Asn117, Asn133 and Asn148) (Norioka *et al.*, 1996). As described above, the five peptides K5a, K5b, K10, K11a and K11b were glycosylated. The K5- and K11-peptides carried two glycosylation sites (Asn60 and Asn74 for the K5-peptide, Asn133 and Asn148 for the K11-peptide), whereas the K10-peptide bore a single site (Asn117) (Table III-6). The sugar compositions of the glycopeptides were assumed on the basis of their deconvoluted masses (Table III-7). In contrast to S_6-RNase, one deoxyhexose is needed to match the calculated masses of these peaks with the observed values. The *N*-glycan of the K10-peptide bears a deoxyhexose and a pentose residue but lacks a hexose residue of the core structure ($HexNAc_2Hex_3$) probably because of a glycosidase existing in the style. If each of the two glycosylation sites in the K5-peptide has the *N*-linked sugar core ($HexNAc_2Hex_3$), the sugar composition must be greater than $HexNAc_4Hex_6$. But the number of hexoses in the K5-peptide was less than four, the K5b-peptide in particular bore no hexose residue, indicative that the glycans in the K5-peptide also undergo digestion by glycosidases. The same phenomenon was found for the K11a- and K11b-peptides (Table III-7). In the case of the K11a-peptide, it was more reasonable to assume that the glycan composed of $HexNAc_2Hex_3deoxyHex_1$ was attached to only one glycosylation site because $HexNAc_2Hex_3$ was the sugar core. At present, it is difficult to determine the sugar composition at each of the two glycosylation sites in the K5- and K11-peptides because of the decomposition of the sugar moieties.

Discussion

Quantitative analysis of the thiol content of a protein generally has depended on the conditions established by Ellman (1959). The problems with that method are that the reaction takes place at about pH 8 and it is reversible. Accordingly, the thiol group contents estimated in terms of liberated nitrothiophenol must be carefully interpreted. S-pyridylethylation of the free thiol groups in a protein at an acidic pH can resolve these problems. Furthermore, the alkylated derivative of free cysteine, PE-Cys, is very stable and easy to identify by amino acid and sequence analyses or mass spectrometry. In addition, this new method can be used at a low-pico mole level. About 95% of the thiol groups in ovalbumin was S-pyridylethylated under optimal conditions (Figure III-1 and Table III-1). More severe reaction conditions (increased 4-VP concentration and longer reaction time) did not improve alkylation yields (Figure III-1). The same tendency was reported by Friedman *et al.* (1970). The reason for this phenomenon is not yet clear. As expected, almost all the thiol groups were blocked in the *N. alata* S_6-RNase. No side reaction by 4-VP was observed, though a certain side reaction at the e-amino group of the lysine residue has been reported to be modified at pH 5.0 (Friedman *et al.*, 1970).

Pepsin or thermolysin at an acidic pH has been used in digestion to obtain cystine-bearing peptide from protein, but digestions with these proteases which have relatively broad substrate specificities are not suitable for obtaining cystine-bearing peptides, in particular when the amount of sample is limited. *N. alata* PE-S_6-RNase and *P. pyrifolia* S_4-RNase therefore were digested with API at pH 6.5 because a similar method was successful for nuclease P1 (Maekawa *et al.*,

1991). As API is strictly specific for lysyl bonds, simple chromatograms were obtained by reversed-phase HPLC (Figures III-3 and III-5). The simple chromatogram provides supporting evidence that no, or little, disulfide exchange took place in the S_6-RNase and S_4-RNase during manipulation. When intact *N. alata* S_6-RNase was digested with API at pH 9.0, a much more complex chromatogram was obtained, possibly because of thiol-disulfide exchange (data not shown). This means that the free thiol groups in S_6-RNase accelerate the disulfide exchange reaction at an alkaline pH. A similar phenomenon has been reported for *N. alata* S_2-RNase bearing the free thiol group (Oxley and Bacic, 1995).

As shown in Figure III-6, the locations of the four disulfide bridges in *N. alata* S_6-RNase are consistent with those in *P. pyrifolia* S_4-RNase, and the eight half-cystine residues forming these bridges are conserved in all the S-RNases of the Solanaceae (Ioerger *et al.*, 1991) and Rosaceae (Norioka *et al.*, 1996) sequenced so far. This suggests that these four disulfide bonds are conserved among all S-RNases and are important in the formation or stabilization of the tertiary structure specific to the S-RNase. In RNase LE, a GSI-unrelated ribonuclease induced by phosphate starvation of cultured cells of *Lycopersicon*, two disulfide bonds have been identified as Cys161-Cys196 and Cys177-Cys188 (Löffeler *et al.*, 1993) which correspond to the two cystine residues near the C-terminus of S_6-RNase or S_4-RNase (Figure III-6). In the fungus RNase T_2-type enzymes, on the one hand the location of the two disulfide linkage in RNase T_2 were chemically determined (Kawata *et al.*, 1989), and on the other hand the five linkages in RNase Rh have been deduced from its tertiary structure (Kurihara *et al.*, 1996) (Figure III-6). Of these linkages, only two, Cys68-Cys118 and Cys191-Cys225 (RNase T_2 numbering), are conserved in the two S-RNases; Cys45-Cys94 and Cys153-Cys182 in S_6-RNase, and Cys48-Cys92 and Cys156-Cys195 in S_4-RNase (Figure III-6). This finding suggests that these

two linkages are essential for the activity of the RNase T_2 family. In fact, the Cys63-Cys112 in RNase Rh (Kurihara *et al.*, 1996) or the Cys45-Cys94 in S_6-RNase (Ishimizu *et al.*, 1995) is, or may be, located near the active site.

The critical difference between the two S-RNases in terms of the half-cystine residues is that two free cysteines are present in *N. alata* S_6-RNase. One, Cys95, is found in many, but not all, of the plant RNase T_2 enzymes (Green, 1994), and the other, Cys77, is conserved in only a few enzymes; *N. alata* S_1-, S_3-RNases and GSI-unrelated RNase X2 of *Petunia inflata* (Green, 1994). Neither residue has been found in the rosaceous S-RNases so far sequenced. These residues therefore do not appear to be directly involved in the catalytic mechanism or the in vivo function of the S-RNase even though S-carboxymethylation of the Cys95 of *N. alata* S_6-RNase causes inactivation (Ishimizu *et al.*, 1995). Further investigation, however, is required to clarify the role(s) of the free cysteine residues.

LC/ESI-MS analyses of the peptides derived from S_6-RNase and S_4-RNase indicate that both bear the *N*-linked glycan with remarkable microheterogeneity and that they undergo no other post-translational modification, such as phosphorylation or *N*-terminal modification. Sugar moiety heterogeneity has been reported for *N. alata* S_1- and S_2-RNases (Oxley and Bacic, 1995). The difference between the *N. alata* S_6- and *P. pyrifolia* S_4-RNase glycans is that deoxypentoses, probably fucoses, are present only in *P. pyrifolia* S_4-RNase and that enzymatic degradation of the glycan by endo-β-mannosidase occurs only in S_4-RNase. Probably, the species and amounts of enzymes that biosynthesize or metabolize glycans differ markedly between *N. alata* and *P. pyrifolia*.

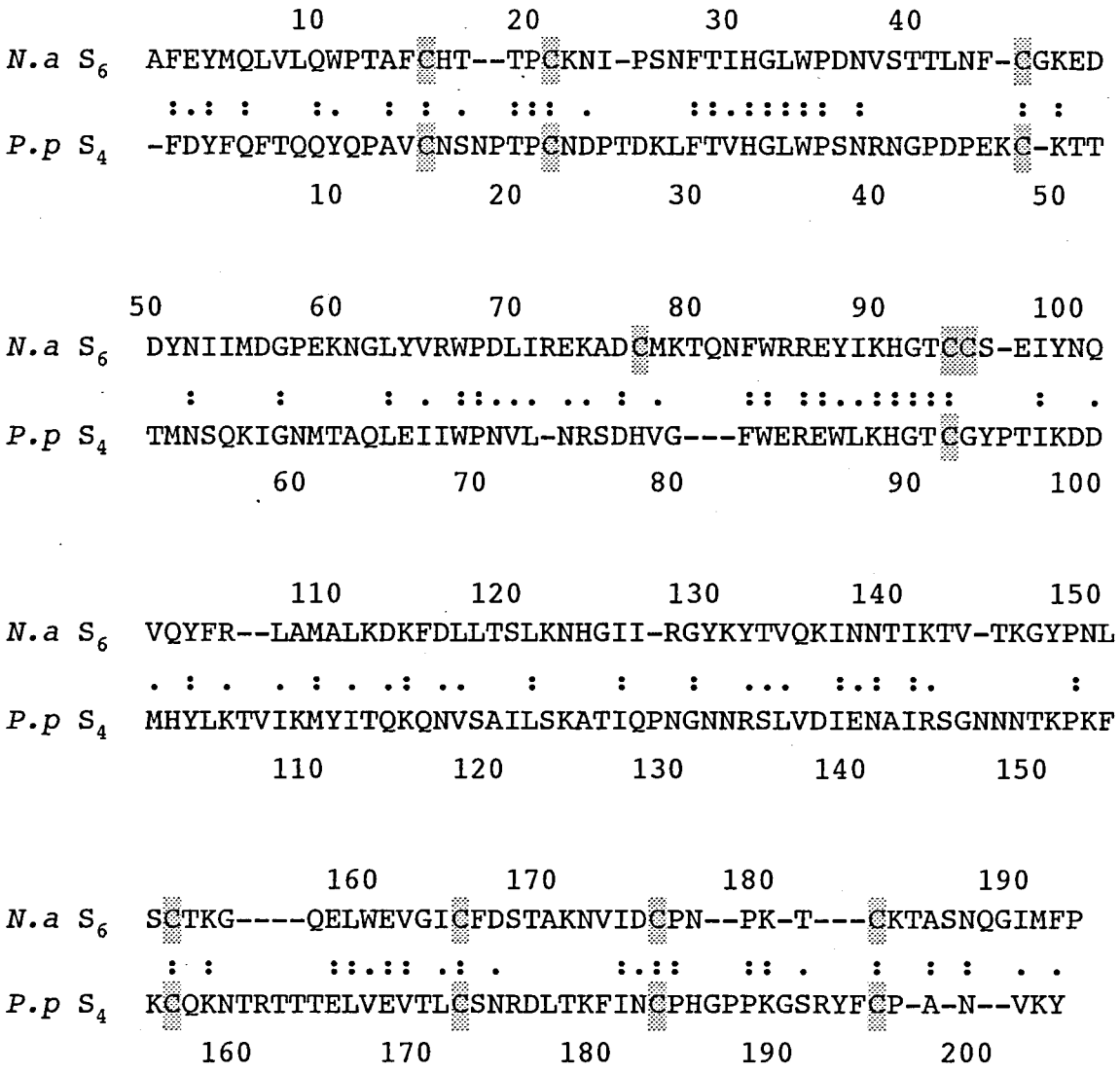


Figure III-1. **Amino acid sequence alignment of *N. alata* S_6-RNase (Anderson *et al.*, 1989) and *P. pyrifolia* S_4-RNase (Norioka *et al.*, 1995).** *N.a* S_6 denotes *N. alata* S_6-RNase, and *P.p* S_4 *P. pyrifolia* S_4-RNase. Amino acid residues identical or similar to each other respectively are designated (:) and (.), the similar amino acid being defined on the basis of T=S, D=N=E=Q, Y=F=W, V=L=M=I and R=K. Half-cystine residues in the two S-RNases are shadowed.

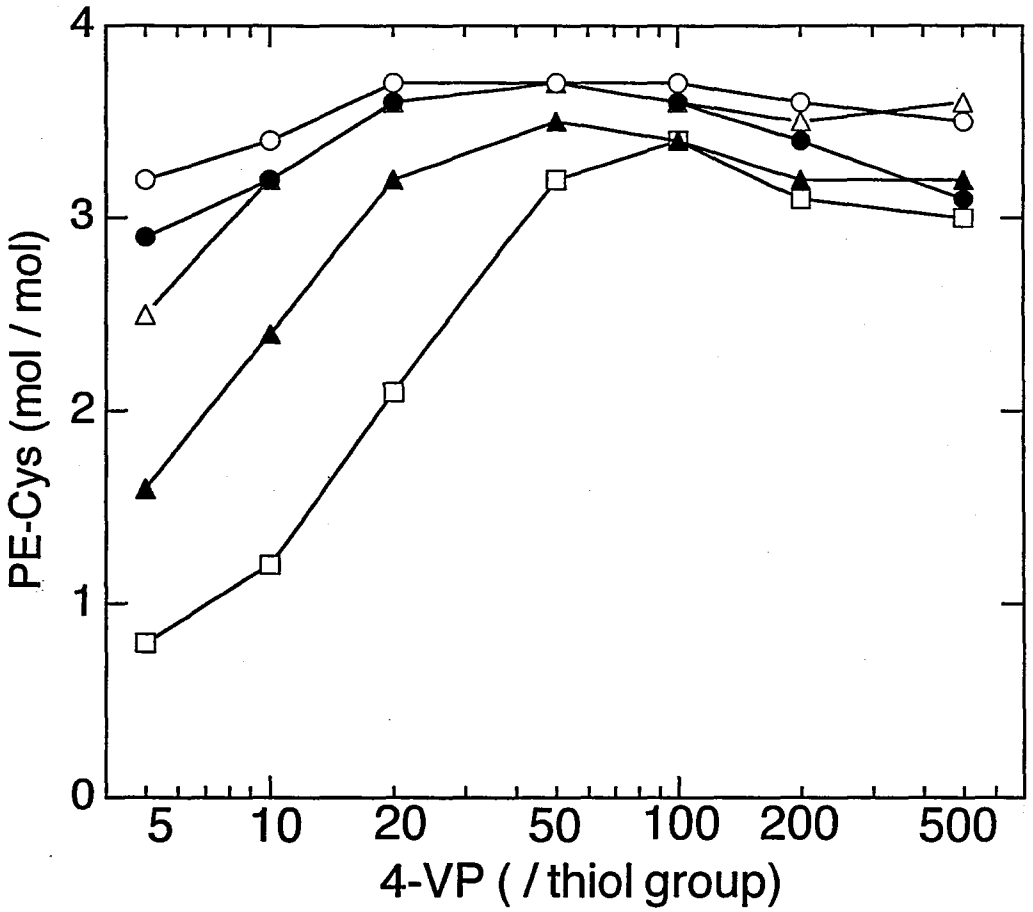


Figure III-2. **Estimation of PE-Cys produced by pyridylethylation of ovalbumin at various pHs.** Ovalubmin was pyridylethylated with various amounts of 4-VP for 4 hr. The PE-Cys produced was estimated by amino acid analysis after reduction and *S*-carboxymethylation. The reaction was carried out at pH 4.5 (□—□), pH 5.0 (▲—▲), pH 5.5 (△—△), pH 6.0 (●—●) and pH 6.5 (○—○).

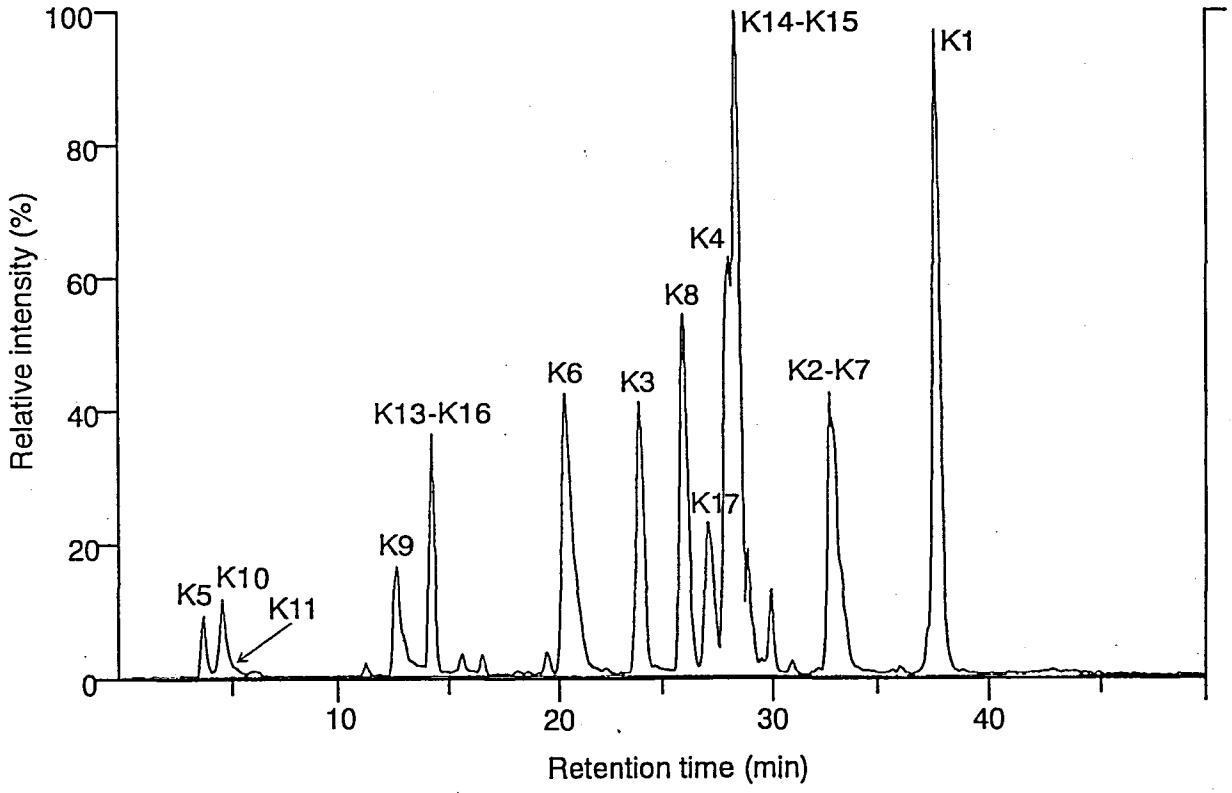


Figure III-3. **LC/ESI-MS-recorded chromatogram for the API-digest of *N. alata* PE-S_6-RNase.** *N. alata* PE-S_6-RNase was digested with API at pH 6.5 then chromatographed on a PLRP-S, 1.0 x 150 mm column (Mishrom Bioresources, Pleasanton, CA, USA). Peptides were monitored by the total ion current. Scans from 330 to 2500 Da were made every 3 seconds.

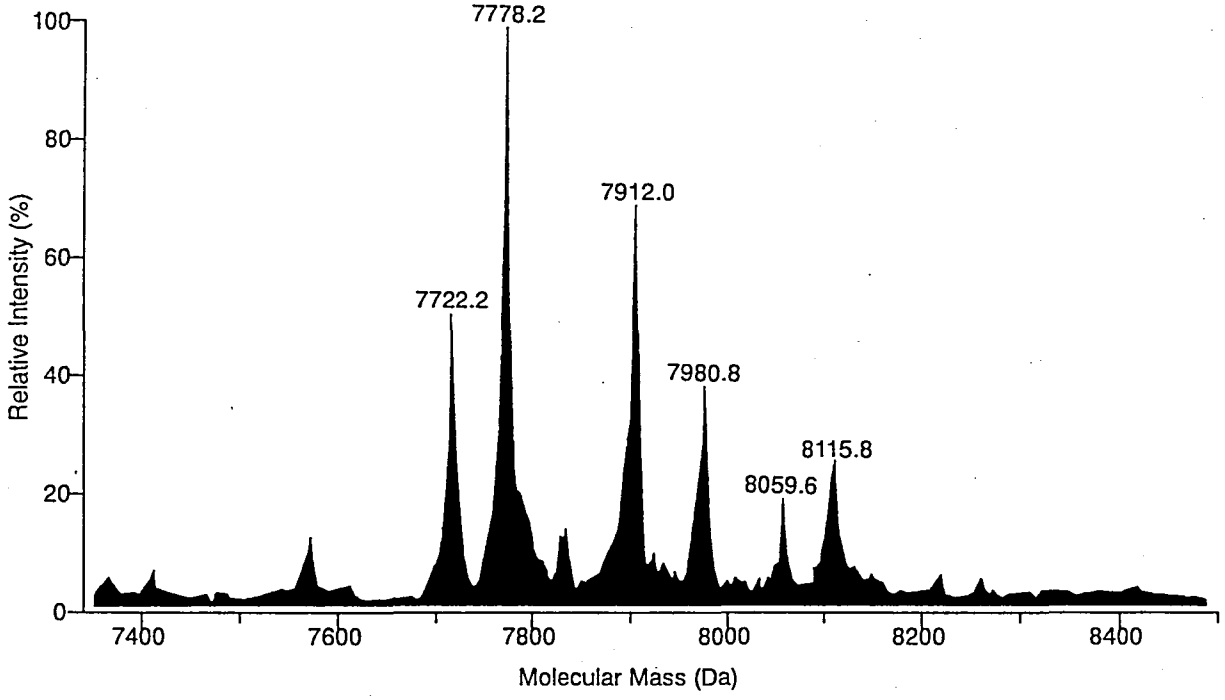


Figure III-4. **Deconvoluted mass spectrum of a glycopeptide from *N. alata* PE-S_6-RNase.** The series of m/z values for the peak eluted at 33.1 min in Figure IV-3. were deconvoluted to molecular masses.

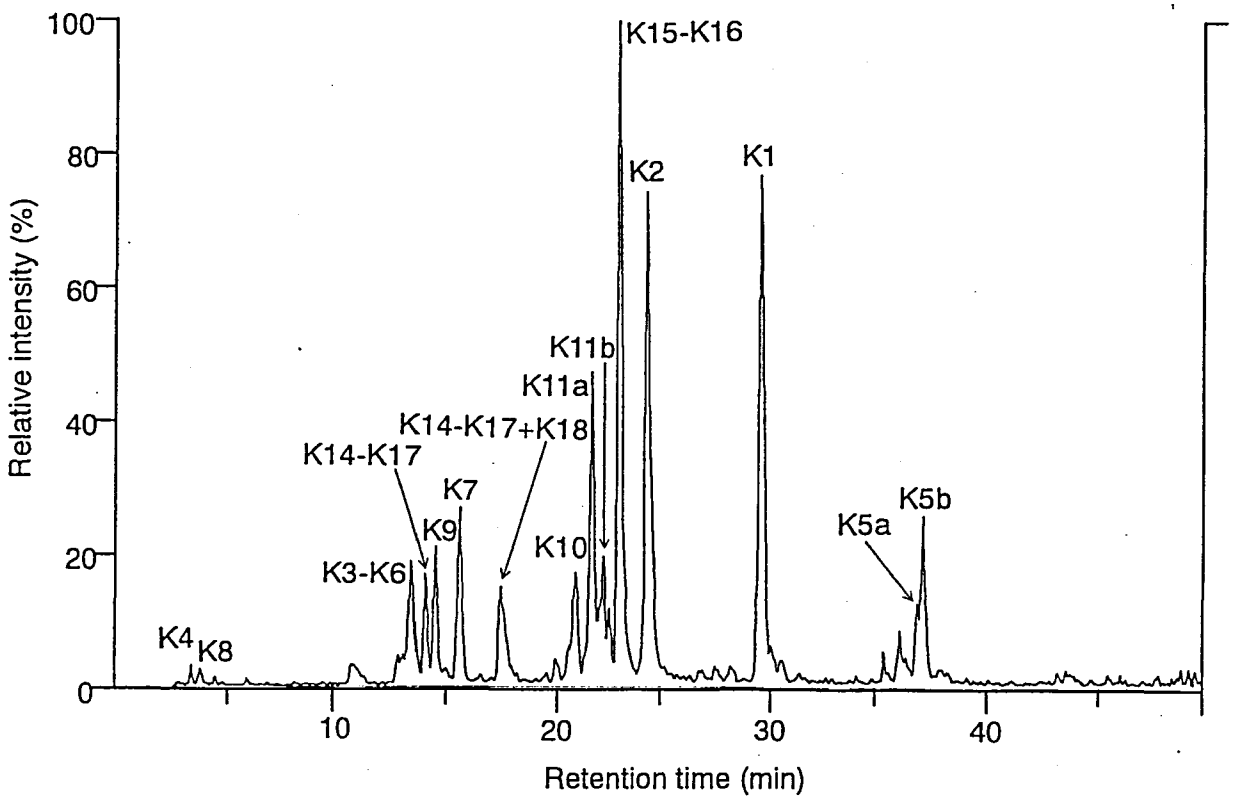


Figure III-5. **LC/ESI-MS-recorded chromatogram for the API-digest of *P. pyrifolia* S_4-RNase.** *P. pyrifolia* S_4-RNase was digested with API at pH 6.5 then chromatographed on a PLRP-S 1.0 x 150 mm column (Mishrom Bioresources, Pleasanton, CA, USA). Peptides were monitored by the total ion current. Scans from 330 to 2500 Da were made every 3 seconds.

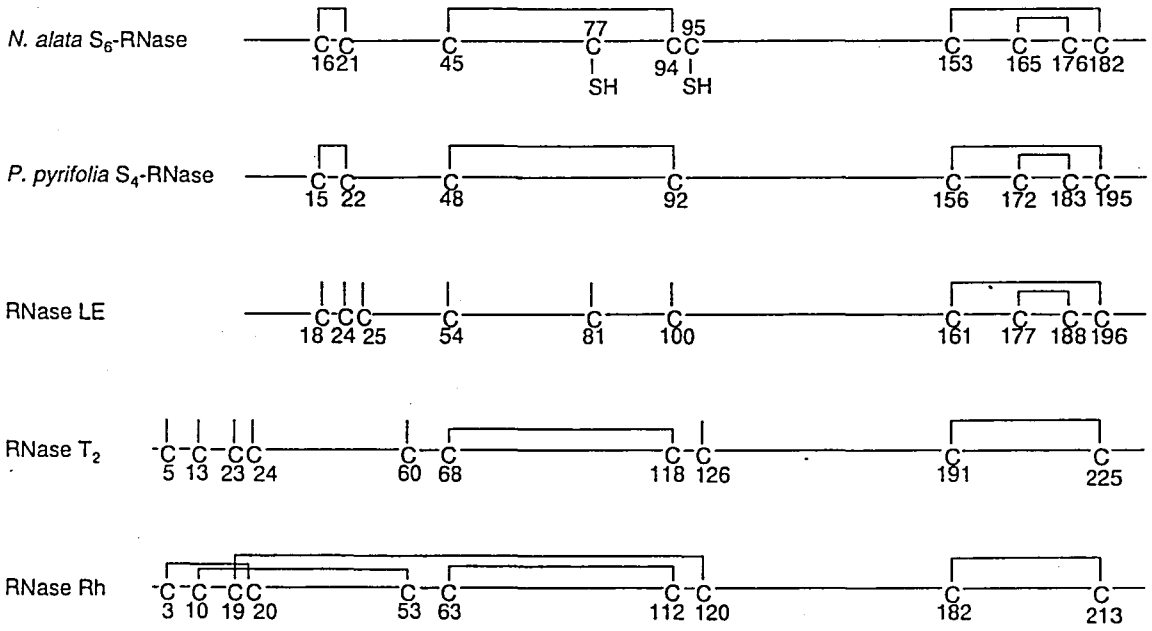


Figure III-6. **Location of free cysteine residues and disulfide bonds in RNase T_2-type enzymes.** Two disulfide bridges were confirmed chemically for RNase LE (Löffeler et al., 1993) and RNase T_2 (Kawata et al., 1989). The location of the disulfide bridges in the fungus RNase Rh was determined from its tertiary structure (Kurihara et al., 1996).

Table III-1. **Amino acid compositions of the RCM-derivative of PE-ovalbumin at various pHs.** *S*-pyridylethylation was done in ovalbumin for 4 hr at various pHs with a 50-fold molar excess of 4-VP to the thiol groups. Values are shown by mol per mol of protein and those in parentheses are based on the amino acid sequence of the native protein (McReynolds *et al.*, 1978).

Amino acid	pH 4.5	pH 5.0	pH 5.5	pH 6.0	pH 6.5
PE-Cys (4)	3.2	3.5	3.7	3.7	3.7
CM-Cys (2)	2.7	2.7	2.4	2.2	2.4
Lys (20)	20.9	20.8	20.7	20.6	20.9
His (7)	7.1	7.1	7.4	7.1	7.5
Trp (3)	1.0[a]	+[a]	+[a]	0[a]	0[a]
Arg (15)	15.2	15.0	14.9	15.7	14.8
Asx (31)	30.9	30.5	30.3	30.3	30.3
Thr (15)	15.5	15.3	15.4	16.7	15.2
Ser (38)	35.8	35.6	35.1	39.0	34.9
Glx (48)	49.4	49.1	48.4	50.1	48.5
Pro (14)	11.1	11.8	10.2	13.7	10.4
Gly (19)	19.3	19.3	19.0	19.2	19.1
Ala (35)	35.0	35.0	34.4	35.0	34.5
Val (31)	28.6[b]	28.7[b]	28.4[b]	28.3[b]	28.6[b]
Met (16)	16.1	15.8	15.9	16.6	15.8
Ile (25)	22.5[b]	22.2[b]	22.6[b]	22.0[b]	22.7[b]
Leu (32)	34.2	34.0	34.3	33.9	34.3
Tyr (10)	10.2	10.2	10.6	10.2	10.6
Phe (20)	19.6	19.5	19.6	20.1	19.6
GlcN	+	+	+	+	+

[a] Low recovery due to decomposition upon HCL hyudrolysis.

[b] Low recovery due to incomplete hydrolysis of the Ile-Ile, Ile-Val and Val-Val linkages.

Table III-2. **Amino acid compositions of the *N. alata* RCM-PE-S_6-RNase and *P. pyrifolia* RCM-PE-S_4-RNase.** Values are shown by mol per mol of protein and those in parentheses are based on the amino acid sequence of the native protein.

Amino acids	*N. alata* RCM-PE-S_6-RNase	*P. pyrifolia* RCM-PE-S_4-RNase
PE-Cys	1.9 } (10)	0.0 } (8)
CM-Cys	7.9	7.7
Lys	16.6 (17)	16.4 (17)
His	3.9 (4)	4.9 (5)
Trp	0.5 (5)[a]	0.0 (4)[a]
Arg	6.1 (6)	8.1 (8)
Asx	23.3 (25)	29.0 (30)
Thr	18.1 (17)	19.6 (20)
Ser	7.4 (7)	9.6 (10)
Glx	16.6 (16)	17.3 (17)
Pro	9.6 (10)	14.0 (15)
Gly	11.1 (11)	10.3 (10)
Ala	7.5 (7)	6.7 (6)
Val	7.6 (8)	9.5 (10)
Met	4.9 (5)	3.6 (4)
Ile	10.8 (14)[b]	10.0 (11)[b]
Leu	13.2 (13)	11.0 (11)
Tyr	8.6 (9)	6.8 (7)
Phe	8.9 (9)	7.6 (8)
GlcNAc	+	+

[a]Low recovery due to decomposition upon HCl hydrolysis.

[b]Low recovery due to incomplete hydrolysis of the Ile-Ile linkage.

Table III-3. **Molecular masses of K-peptides from *N. alata* PE-S_6-RNase.** The average masses were calculated from the amino acid sequence of S_6-RNase. K-peptides obtained by API digestion are numbered in order from the *N*-terminus of the protein.

Peptide No.	Position	Retention time (min)	Molecular mass (Da)		Sequence
			Calculated	Observed	
K1[b]	1-22	37.5	2613.1	2613.7	16 21 AFEYMQLVLQWPTAFCHTTPCK
K2-K7[a, b, c]	23-47 91-112	33.1	see Table 5		CHO CHO 45 NIPSNFTIHGLWPDNVSTTLNFCGK 95 HGTCCSEIYNQVQYFRLAMALK 94 PE
K3	48-60	23.6	1538.7	1538.1	EDDYNIIMDGPEK
K4	61-74	27.8	1759.1	1758.5	NGLYVRWPDLIREK
K5[a]	75-79	3.6	671.8	671.5	77 ADCMK PE
K6	80-90	20.1	1540.8	1540.6	TQNFWRREYIK
K8	113-122	25.6	1179.4	1179.4	DKFDLLTSLK
K9	123-131	12.3	1057.2	1057.0	NHGIIRGYK
K10	132-136	4.4	637.7	637.4	YTVQK
K11[c]	137-142	4.9	see Table 5		CHO INNTIK
K12	143-146	—[d]	448.5		TVTK
K13-K16[b]	147-155 181-183	13.9	1330.5	1330.2	153 GYPNLSCTK TCK 182
K14-K15[b]	156-171 172-180	28.3	2780.1	2780.6	165 GQELWEVGICFDSTAK NVIDCPNPK 176
K17	184-193	26.8	1065.2	1064.9	TASNQGIMFP

[a] the fragment bears a pyridylethylcysteine, which adds 105.1Da to the mass.
[b] the fragment has a disulfide bridge, which subtracts 2.0 Da from the mass.
[c] the fragment has *N*-glycan(s), which adds appropriate mass to the mass.
[d] not identified as the fragment was eluted in the solvent front.

Table III-4. **Amino acid sequence analysis of the peak eluted at 33.1 min in Figure IV-3.** This peak was collected and analyzed with a gas-phase protein sequencer.

Cycle	Amounts of PTH-amino acid (pmol)	
1	Asn (16.6)	His (3.2)
2	Ile (30.0)	Gly (22.2)
3	Pro (23.2)	Thr (6.2)
4	Ser (21.2)	dehydro Ala (21.2)
5	-[a]	PE-Cys (14.9)
6	Phe (18.9)	Ser (3.6)

[a] PTH-asparagine was not extracted by butyl chloride due to attachment of an *N*-glycan.

Table III-5. **Sugar compositions of *N*-linked glycopeptides from *N. alata* S_6-RNase.** Molecular masses are calculated from the amino acid sequence and putative sugar composition. The K2-K7-peptide has two *N*-glycosylation sites (Asn27 and Asn37), and the K11-peptide has one glycosylation site (Asn138).

Peptide No.	Position	Retention time (min)	Molecular mass (Da)		Expected sugar compasition
			Calculated	Observed	
K2-K7[a]	23-47 91-112	33.1	7723.7	7722.2	$HexNAc_4Hex_9$
			7775.8	7778.2	$HexNAc_6Hex_6Pen_1$
			7907.9	7912.0	$HexNAc_6Hex_6Pen_2$
			7979.0	7980.8	$HexNAc_7Hex_6Pen_1$
			8059.0	8059.6	$HexNAc_5Hex_9Pen_1$
			8111.1	8115.8	$HexNAc_7Hex_6Pen_2$
K11	137-142	4.9	1929.9	1929.8	$HexNAc_3Hex_3Pen_1$
			2092.1	2091.7	$HexNAc_3Hex_4Pen_1$

[a] the fragment has a pyridylethylcysteine which adds 105.1 Da to the mass and a disulfide bridge, which subtracts 2.0 Da from it.

Table III-6. **Molecular masses of the K-peptides from *P. pyrifolia* S_4-RNase.** Average masses calculated are based on the amino acid sequence of S_4-RNase. K-peptides obtained by API digestion are numbered in order from the the *N*-terminus of S_4-RNase.

Peptide No.	Position	Retention time (min)	Molecular mass (Da)		Sequence
			Calculated	Observed	
K1[a]	1-28	29.5	3267.5	3268.2	FDYFQFTQQYQPAVCNSNPTPCNDPTDK (15–22)
K2	29-47	24.2	2164.4	2164.6	LFTVHGLWPSNRNGPDPEK
K3-K6[a]	48-49 89-98	13.3	1323.6	1323.0	CK (48) HGTCGYPTIK
K4	50-57	3.5	910.0	909.6	TTTMNSQK (92)
K5[c]	58-88	36.8 37.1	see Table 7		IGNMTAQLEIIWPNVLNRSDHVGFWEREWLK (CHO) (CHO)
K7	99-105	15.4	921.1	920.6	DDMHYLK
K8	106-109	3.8	459.6	459.5	TVIK
K9	110-115	14.4	783.0	782.5	MYITQK
K10[c]	116-124	20.5 20.9	see Table 7		QNVSAILSK (CHO)
K11[c]	125-151	21.6 22.1	see Table 7		ATIQPNGNNRSLVDIENAIRSGNNNTK (CHO) (CHO)
K12[b]	152-153	—	243.3		PK
K13[b]	154-155	—	293.4		FK
K14-K17[a]	156-158 190-200	14.0	1616.9	1616.5	CQK (156) GSRYFCPANVK (195)
K14-K17+K18[a]	156-158 190-201	17.3	1780.1	1779.7	CQK (156) GSRYFCPANVKY (195)
K15-K16[a]	159-179 180-189	23.0	3502.0	3502.0	NTRTTTELVEVTLCSNRDLTK (172) FINCPHGPPK (183)
K18[b]	201	—	181.2		Y

[a] the fragment has a disulfide bridge which subtracts 2.0 Da from the mass.

[b] the fragment was not detected because its molecular mass was less than the minimum scan range limit.

[c] the fragment bears an *N*-glycan(s), which adds an appropriate mass to the mass.

Table III-7. **Sugar compositions of the *N*-linked glycopeptides from *P. pyrifolia* S_4-RNase.** Molecular masses calculated are based on the amino acid sequence and putative sugar composition. Each of K5- and K11-peptides has two *N*-glycosylation sites (Asn60 and Asn74 for K5-peptide, Asn133 and Asn148 for K11-peptide), and K10-peptide has one glycosylation site (Asn117).

Peptide No.	Position	Retention time (min)	Molecular mass (Da)		Expected sugar composition
			Calculated	Observed	
K5a	58-88	36.8	5168.6	5168.7	$HexNAc_4Hex_2deoxyHex_1Pen_1$
			5330.8	5330.0	$HexNAc_4Hex_3deoxyHex_1Pen_1$
			5534.0	5532.8	$HexNAc_5Hex_3deoxyHex_1Pen_1$
K5b		37.1	4362.9	4363.0	$HexNAc_3$
			4566.1	4566.3	$HexNAc_4$
K10	116-124	20.9	1835.9	1834.2	$HexNAc_2Hex_2deoxyHex_1$
			1968.1	1966.3	$HexNAc_2Hex_2deoxyHex_1Pen_1$
			2171.2	2170.2	$HexNAc_3Hex_2deoxyHex_1Pen_1$
K11a	125-151	21.6	3950.1	3949.4	$HexNAc_2Hex_3deoxyHex_1$
K11b		22.1	3723.9	3723.7	$HexNAc_4$

Chapter IV
Structures of *N*-glycans in *P. pyrifolia* S-RNases

Introduction

All S-RNases have Asn-Xaa-Ser / Thr consensus sequence for the site of potential *N*-glycosilation (Norioka *et al.*, 1995, Chapter II, Ishimizu *et al.*, 1998b). The molecular weight of purified *P. pyrifolia* S-RNases observed by MALDI / TOF-MS are 2000 or more higher than those calculated from the primary structures (unpublished data of our laboratory), and *P. pyrifolia* S-RNase was reacted with concanavalin A and wheat germ agglutinin (Sassa *et al.*, 1993), meaning that S-RNase actually has glycans. Since the *N*-glycosylation sites are variable among S-RNase, this diversity may be responsible for discrimination between self and non-self pollen.

To date the structures of *N*-glycan enzymatically released from *Nicotiana alata* S-RNases were clarified (Oxley and Bacic, 1995; Oxley *et al.*, 1996). In the *P. pyrifolia* S-RNases, two to five potential *N*-glycosylation sites were observed in each amino acid sequences (Ishimizu *et al.*, 1998b). Liquid chromatography / electrospray ionization-mass spectrometry (LC/ESI-MS) analysis of *Achromobacter* protease I (API)-digest of the *P. pyrifolia* S_4-RNase showed microheterogeneity of *N*-glycan (Ishimizu *et al.*, 1996b). As for the functional analysis of carbohydrate moiety of S-RNase, the experiment using the transgenic plants demonstrated that *N*-glycan of the *Petunia inflata* S_3-RNase, which was attached to the conserved site in solanaceous S-RNases, is not required for SI reaction (Karunanandaa *et al.*, 1994).

In this study, structures of *N*-glycans attached to *P. pyrifolia* S_4-RNase were determined using high sensitive methods. Chemically released *N*-glycans were pyridylaminated (Hase *et al.*, 1994) and partially hydrolyzed. They were analyzed by two kinds of chromatographies (two-dimensional sugar map) (Hase *et al.*, 1988). Identification and quantitative analysis of glycans at each *N*-glycosylation site were analyzed by LC/ESI-MS analysis of protease-digest of S_4-RNase. To discuss the role of *N*-glycan of S-RNase in SI interaction, *N*-glycans of other six *P. pyrifolia* S-RNases were also analyzed by LC/ESI-MS. From the results obtained by these analyses, the function of *N*-glycan at each site and biosynthesis pathway of these glycans of *P. pyrifolia* S-RNases will be discussed.

Materials and Methods

Materials. S-RNases were purified from styles of *P. pyrifolia*, as reported elsewhere (Ishimizu *et al.*, 1996b). S_1- and S_6-RNases were purified from styles of the cultivar 'Imamuraaki' (S_1S_6), S_2- and S_4-RNases from 'Nijisseiki' (S_2S_4), S_3- and S_5-RNases from 'Hosui' (S_3S_5), and S_7-RNase from 'Okusankichi' (S_5S_7). α-Mannosidase (Jack bean) and α-L-fucosidase (*Charonia lampas*) were purchased from Seikagaku Kogyo, *N*-acetyl-β-D-glucosaminidase (*Diplococcus pneumoniae*) from Boehringer Mannheim. PA-isomaltooligosaccharides were purchased from Takara Biomedicals. The structures and abbreviations of authentic PA-sugar chains used in this study were listed in Table IV-1. GN and GN2 (Makino *et al.*, 1996); MN (Suzuki *et al.*, 1991); M1, M2B, M3B, M4C, and M5A (OKu *et al.*, 1990); and MX, M2X, M3X, M4X, M2FX, and M3FX (Kimura *et*

al., 1987) were prepared as described previously. AG1 was purchased from Takara Biomedicals.

Preparation of PA-sugar chains from S_4-RNase. Sugar chains were released from 3 mg of S_4-RNase by hydrazinolysis and the free amino groups of the hydrazinolyzate were *N*-acetylated. The released sugar chains were pyridylaminated as described previously (Kuraya and Hase, 1992) and separated by reversed-phase HPLC with Cosmosil 5C18P column (4.6 mm x 150 mm) (Nacalai tesque). The mixture of PA-sugar chains was applied to the column equilibrated with 0.025 % n-buthanol in 20 mM ammonium acetate, pH 4.0. The gradient of n-buthanol was developed linealy to 0.5 % for 55 min at a flow rate of 1.5 ml / min. PA-sugar chains were detected by their fluorescence (excitation wavelength, 320 nm; emission wavelength, 400 nm). The peaks containing more than two kinds of PA-sugar chains were collected and rechromatographed on a NH_2P column (4.0 mm x 50 mm) (Asahikasei). Elution and detection carried out under the same condition as described previously (Makino *et al.*, 1996).

Reducing-end analysis of PA-sugar chains. PA-sugar chains were hydrolyzed with 4 M hydrochloride at 100 °C for 8 h. The hydrolyzates were applied to a TSKgel Sugar AXI column (4.6 mm x 150 mm) (Tosoh) after *N*-acetylation (Hase *et al.*, 1992).

Mono- and di-saccharide analysis. PA-mono- and di-saccharides were analyzed with a TSKgel Sugar AXI column (4.6 mm x 150 mm) (Tosoh) after *N*-acetylation (Hase *et al.*, 1992).

Exoglycosidase digestion of PA-sugar chains. Digestion with α-mannosidase (0.2 U) was performed in 30 μl of 40 mM sodium citrate buffer, pH 4.5, at 37 °C for 16 hr; α-fucosidase (0.05 U) in 25 μl of 0.1 M sodium acetate buffer containing 0.5 M sodium chloride, pH 3.9, at 37 °C for 24 hr; *N*-acetyl-β-D-glucosaminidase (5 mU) in 25 μl 50 mM sodium citrate buffer, pH 5.3, at 37 °C for 24 hr. Each reaction was stopped by boiling the solution for 2 min and an aliquot of the digest was analyzed by HPLC.

Partial acid hydrolysis of PA-sugar chains. PA-sugar chains were hydrolyzed with 100 μl of 1 M trifluoroacetic acid at 100 °C for 7 to 20 min. The reducing-end terminal PA-sugar chains were collected as described previously (Makino *et al.*, 1996) and analyzed by HPLCs.

High performance liquid chromatography of PA-sugar chains. Reversed-phase chromatography of analyzing PA-sugar chains was performed as described previously (Yanagida *et al.*, 1998). The reversed-phase scale (RPS) of each PA-sugar chain was introduced according to the method of Yanagida *et al.* (1998).

Size-fractionation chromatography was performed on a NH_2P column (2 mm x 70 mm) (Showa Denko). Elution and detection of PA-sugar chains were carried out under the same conditions as described previously (Makino *et al.*, 1996). Molecular size of each PA-sugar chain was shown in terms of a glucose unit based on the elution times of PA-isomaltooligosaccharides.

The HPLC profiles (RPSs and molecular sizes) of authentic PA-sugar chains were observed as shown in Table IV-1. Those of PA-sugar chains and their partial hydrolyzates were compared with those of authentic PA-sugar chains.

Reduction and *S*-carboxymethylation of S-RNase. Reduction and *S*-carboxymethylation of S-RNases were done by the method of Crestfield *et al.* (1963). The reduced and *S*-carboxymethylated (RCM-) protein was desalted in a Fast Desalting column (3.2 x 100 mm) equilibrated with 0.05% trifluoroacetic acid using SMART system (Pharmacia).

LC/ESI-MS analysis of protease digests of S-RNases. RCM-S_4-RNase was digested with *Staphylococcus aureus* V8 protease (V8) in 10 mM ammonium bicarbonate buffer, pH 7.8, at 37 °C for 12 hr at the enzyme / substrate ratio of 1 / 50 (mol / mol). After boiling the mixture for 2 min, subsequent digestion with *Achromobacter protease* I (API) in the same buffer at 37°C for 4 hr at the enzyme / substrate ratio of 1 / 100. The other six RCM-S-RNases (*P. pyrifolia* S_1-, S_2-, S_3-, S_5-, S_6-, and S_7-RNases) were digested with API in 10 mM Tris-HCl buffer, pH 9.0, at 37 °C for 4 hr at the enzyme / substrate ratio of 1 / 200. The digests were analyzed on a Perkin-Elmer Sciex API-III triple quadrupole mass spectrometer equipped with an electrospray ionization system to which a 10 nm PLRP-S column (1 x 150 mm) (Michrom Bioresource) is connected via an interface operating in the electrospray ionization mode with 5 kV of needle voltage. Each digest was applied to the column equilibrated with 0.1 % formic acid in 95% water and 5% acetonitrile. The gradient of organic solvent (0.1% formic acid in acetonitrile) was developed linearly to 65% for 60 min at a flow rate of 40 μl / min. Each peptide eluted was injected to the mass spectrometer via the interface. The quadrupole was scanned over 300-2400 Da every 4.2 seconds. The series of m / z values of the multiple charged peptides in each peak were deconvoluted to a given molecular mass. The mass values were assigned to the peptides expected to produce by digestion with API, or V8 and API based on the amino acid sequences of S-RNases (Norioka *et al.*, 1995; Ishimizu *et al.*, 1998a).

Results

Separation of PA-sugar chains from S_4-RNase. Anion-exchange HPLC analysis of PA-sugar chains obtained from S_4-RNase detected no acidic sugar chain (data not shown). PA-sugar chains were fractionated by reversed-phase HPLC (Figure IV-1(a)). Total yield of PA-sugar chains calculated from the area of peaks in this chromatogram was 25 %. The fractions containing more than two kinds of PA-sugar chains were then separated by size-fractionation HPLC (Figure IV-1 (b), (c), and (d)). Finally, twelve PA-sugar chains were isolated. PA-sugar chains of which yield were less than 100 pmol were neglected.

Reducing-end analysis of PA-sugar chains. PA-monosaccharide liberated from each PA-sugar chain by acid hydrolysis was listed in Table IV-2. All PA-sugar chains except the fraction 2 had an *N*-acetylglucosamine as reducing-end residue. PA-*N*-acetylmannosamine was detected for the fraction 2 after acid hydrolysis.

Structural analysis of PA-sugar chains. Sequential exoglycosidase digestions, partial acid hydrolyses, and two kinds of HPLCs were carried out for the PA-sugar chains to identify their structures. The profiles for each PA-sugar chain (Table IV-2) were compared with those of authentic PA-sugar chains listed in Table IV-1. Detailed procedures for the identification of the structures of 12 PA-sugar chains are described below.

Fraction 1—The sugar chain of this fraction was assigned to GN from RPS, molecular size, and monosaccharide analysis of this PA-sugar chain.

Fraction 2—This PA-sugar chain had *N*-acetylmannosamine residue as a reducing-end (Table IV-2) and appeared to be PA-disaccharide judging from its molecular size. The digestion with *N*-acetyl-β-D-glucosaminidase of this fraction produced a PA-sugar chain of which chromatographic behavior was similar to that of PA-*N*-acetylmannosamine. These results indicate that fraction 2 is PA-*N*-acetylglucosaminyl *N*-acetylmannosamine, which is probably produced by epimerization of GN2 (fraction 3) during derivatization of PA-sugar chains.

Fraction 3—This fraction was eluted at the same position as GN2 on the reversed-phase, size-fractionation, and sugar AX-I column chromatograms (Table IV-2). Digestion of this fraction with *N*-acetyl-β-D-glucosaminidase produced PA-sugar chain eluted at the same position as GN.

Fraction 4-1—RPS and molecular size of this fraction were similar to those of M3FX (Table IV-2). Partial acid hydrolysis of this fraction produced the PA-sugar chain corresponding to M3B. The digest of this fraction with α-mannosidase was eluted at the same position as MFX. The product of digestion with α-L-fucosidase was eluted at the same position as MX. The partial acid hydrolysis of this product gave four PA-sugar chains of which the elution times on two chromatographies were consistent with those of MX, M1, GN2, and GN, respectively. These analyses indicate that fraction 4-1 is M3FX.

Fraction 4-2—The profiles of chromatographic behavior of fraction 4-2 corresponded to the RPS and molecular size of AG1FX (Tables IV-1 and 2). Partial acid hydorolyzate of this fraction contained the PA-sugar chains eluted at the same position as AG1 in the two kinds of HPLCs. The digest of this fraction with *N*-acetyl-β-D-glucosaminidase was eluted at the same position as M3FX. The following sequential exoglycosidase digestion and partial acid hydrolysis

gave the same results as those obtained for fraction 4-1. These results suggest that fraction 4-2 is AG1FX.

Fraction 5—The profiles of chromatographic behavior for this fraction were the same as M2FX (Table IV-2). Partial acid hydrolysis of this fraction produced the PA-sugar chain corresponding to M2B. The digest of this fraction with α-mannosidase was eluted at the same position as MFX. Structural analysis of this product was performed with α-L-fucosidase digestion and partial acid hydrolysis as described in fraction 4-1 and the same results were obtained as fraction 4-1. These results indicate that fraction 5 is M2FX.

Fraction 6-1—Fraction 6-1 and its digest with α-mannosidase were eluted at the same positions as M3B and M1, respectively (Table IV-2). The following partial acid hydrolysis gave three PA-sugar chains. The eluted positions of the three PA-sugar chains in HPLCs corresponded to those of M1, GN2, and GN. These results indicate that fraction 6-1 is M3B.

Fraction 6-2—Fraction 6-2 was eluted at the same position as M3X (Table IV-2). Partial acid hydorolyzate of this fraction contained the PA-sugar chains eluted at the same position as M3B in the two kinds of HPLCs. The digest of this fraction with α-mannosidase was eluted at the same position as MX. Analysis of this digest with partial acid hydrolysis gave the same results as those obtained for fraction 4-1. These chromatographic profiles indicate that fraction 6-2 is M3X.

Fraction 6-3—This fraction was eluted at the same position as AG1X (Tables IV-1 and 2). Partial acid hydrolysis of this fraction produced the PA-sugar chain eluted at the same position as AG1. The digest of this fraction with *N*-acetyl-β-D-glucosaminidase was eluted at the same position as M3X. The following sequential exoglycosidase digestion and partial acid hydrolysis gave the same results as those obtained for fraction 6-2. These results indicate that fraction 6-3 is AG1X.

Fraction 7-1—This fraction was eluted in the HPLCs at the same position as AG1 (Table IV-2). After this fraction was digested with *N*-acetyl-β-D-glucosaminidase, the new peak appeared at the same position as M3B. The following sequential exoglycosidase digestion and partial acid hydrolysis gave the same results as those of fraction 6-1. These results indicate that fraction 7-1 is AG1.

Fraction 7-2—This fraction was eluted at the same position as M4X (Table IV-2). Partial acid hydrolysis of this fraction produced the PA-sugar chain eluted at the same position as M4C. The digest of this fraction with α-mannosidase was eluted at the same position as M1. The following analysis gave the same results as those of fraction 4-1. These results suggest that fraction 7-2 is M4X.

Fraction 7-3—This fraction was eluted at the same position as M5A (Table IV-2). The profiles of the HPLCs of the digest with α-mannosidase and its acid hydrolyzate indicate that fraction 7-3 is M5A.

The structures of 12 PA-sugar chains derivatized from S_4-RNase were proposed as shown in Table IV-3. The peak area on the chromatogram of Figure IV-1 (a) revealed that more than 70 % of glycans were the short type glycans (fractions 1, 2, and 3 in Figure IV-1).

Glycans attached to each *N*-glycosylation site of S_4-RNase. S_4-RNase has five potential *N*-glycosylation sites, Asn 60, Asn 74, Asn 117, Asn 133, and Asn 148 (Figure IV-2). To identify glycans attached to each *N*-glycosylation, mass spectrometrical analysis of the protease digest of S_4-RNase was carried out. Two proteases, V8 and API, were used for fragmentation of S_4-RNase to generate glycopeptides bearing one *N*-glycosylation site. Peptide mapping of the V8 and API digest of S_4-RNase was constructed by LC/ESI-MS analysis (Figure IV-3). Almost all the peaks were assigned to peptides expected to produce by digestion with two proteases. In these peaks, seven peaks marked

with asterisks in Figure IV-3 were assigned to glycopeptides because these peaks were able to be assigned to molecular masses calculate for the possible peptides bearing an appropriate glycans. Their mass values and expected sugar composition were listed in Table IV-4.

The peak eluted at 19.4 min gave two deconvoluted mass values because of the heterogeneity of the sugar moiety. When one of the two deconvoluted peaks (1179.1 Da) was tentatively assigned to the glycopeptide from Ile58 to Glu66 bearing one *N*-acetylhexosamine residue, the observed mass was consistent with the calculated mass (1179.3 Da). The other peak (1382.0 Da) was presumed to be the same peptide bearing two *N*-acetylhexosamine residues. Therefore, the sugar chains attached to Asn 60 were presumed to be an *N*-acetylglucosamine and a chitobiose, which correspond to GN and GN2 in Table IV-3, respectively.

The peak at 29.7 min was assigned to the glycopeptide from Ile67 to Trp83 bearing an *N*-acetylglucosamine (2285.0 Da) or a chitobiose (2488.4 Da), which corresponds to GN and GN2 in Table IV-3, respectively. The peak at 29.2 min gave three deconvoluted mass values, 3091.6 Da, 3253.2 Da, and 3457.3 Da. The mass values of 3091.6 Da corresponded to the calculated mass value of a glycopeptide Ile 67-Trp 83 bearing two *N*-acetylhexosamine, two hexoses, one deoxyhexose, and one pentose residues ($HexNAc_2Hex_2deoxyHex_1Pen_1$) (3091.3 Da). The mass values of 3253.2 and 3457.3 Da corresponded to the 3091.6 Da peptide plus hexose ($HexNAc_2Hex_3deoxyHex_1Pen_1$) and an additional *N*-acetylhexosamine ($HexNAc_3Hex_3deoxyHex_1Pen_1$), respectively. The mass of peak at 27.2 min corresponded to that of a glycopeptide Ile 67-Glu 85 bearing two HexNAc, but the ion intensity of this peak was much lower than those of peaks at 29.2 min and 29.7 min. These results indicate that a chitobiose (GN in Table IV-3) is a major *N*-glycan attached to Asn 74, and *N*-

acetylglucosamine (GN2) and complex type glycan with xylose and fucose residues (M2FX, M3FX, and AG1FX) are minor components.

Eight major deconvoluted mass peaks were observed for the peak at 16.0 min. The mass value of 1689.2 Da corresponded to the calculated mass of a glycopeptide Gln116-Lys124 bearing two *N*-acetylhexosamine and two hexose residues ($HexNAc_2Hex_2$). The mass value of 1851.2 Da corresponded to a glycopeptide bearing $HexNAc_2Hex_3$, and the 1984.0 Da peak to a glycopeptide with $HexNAc_2Hex_3Pen_1$. The 2014.0, 2054.4, 2146.0, 2157.6, 2168.8 Da peaks corresponded to the calculated masses of glycopeptides bearing $HexNAc_2Hex_4$, $HexNAc_3Hex_3$, $HexNAc_2Hex_4Pen_1$, $HexNAc_2Hex_5$, $HexNAc_3Hex_3Pen_1$. These results indicate that high mannose and complex type sugar chains with or without xylose residue (including M3B, M3X, AG1X, M4C, M4X, and M5A in Table IV-3) are attached to Asn117.

In the peak eluted at 21.3 min, three deconvoluted mass values were observed. The 1740.8 Da peak corresponded to the calculated mass of the peptide Ala125-Glu140 containing Asn133. The 1944.4 and 2147.6 Da peaks corresponded respectively to the 1740.8 Da peptide plus one and two *N*-acetylhexosamine residue(s). Therefore, there were three forms of *N*-glycans, nonglycosylated, one *N*-acetylglucosamine and a chitobiose (GN and GN2 in Table IV-3) at the site Asn 133.

In the peak eluted at 3.8 min, 1594.5 Da peak was observed. This observed mass was consistent with the calculated mass of a glycopeptide Asn141-Lys151 bearing two *N*-actylhexosamine. No additional mass was observed. These results indicate that only a chitobiose (GN2 in Table IV-3) is attached to Asn 148.

The quantitative ratio of each sugar chains estimated from their relative ion intensities was listed in Table IV-3. Since these values roughly corresponded to those calculated from the peak area on the reversed-phase chromatogram

(Table IV-3), the ratio of sugar chains at each glycosylation site listed in Tables IV-4 and 5 were reliable.

Comparison of the structures and positions of *N*-glycans among seven *P. Pyrifolia* S-RNases. The numbers of potential *N*-glycosylation sites of *P. pyrifolia* S-RNases are two (S_3- and S_5-RNases) to five (S_4-, S_6-, and S_7-RNases) (Figure IV-2). The location of the *N*-glycosylation sites are different from each other. To discuss which part of S-RNases is involved in *S*-allele-specific recognition, the structures and positions of *N*-glycans of other six *P. pyrifolia* S-RNases were analyzed by LC/ESI-MS of their API digests, and the structures and positions of *N*-glycans among seven S-RNases were compared with one another. Since the developmental and spatial expression pattern of each S-RNase are same as that of S_4-RNase, it is able to suppose that other six S-RNases have the *N*-glycans of which structures are similar to those of S_4-RNase.

Sugar compositions of the glycopeptides from the API-digests of the six S-RNases were presumed as listed in Table IV-5. Actually, all the potential *N*-glycosylation sites of S-RNases were glycosylated. In these glycopeptides, there were four peptides containing two *N*-glycosylation sites; a glycopeptide containing Asn 60 and Asn 74 in S_1-RNase, a glycopeptide containing Asn 130 and Asn 148 in S_1-RNase, a glycopeptide Asn 161 and Asn 176 in S_6-RNase, and a glycopeptide containing Asn 47 and Asn 50 in S_7-RNase. The latter three peptides were presumed to have $HexNAc_2$, $HexNAc_3$, $HexNAc_4$ as glycans from the observed molecular mass. Therefore, the short type glycans probably are linked to these asparagine residues. Asn 60 and Asn 74 are conserved between S_1- and S_4-RNases. The amino acid sequence around these residues of S_1-RNase are the same as that of S_4-RNase. Since the structure of glycan is affected by its local primary and secondary structures, *N*-glycans of Asn 60 and Asn 74 appear to be similar between S_1- and S_4-RNases. On the basis of this

hypothesis, Asn 60 of S_1-RNase appeared to have only the short type glycans and Asn 74 did the short type and the complex type sugar chain with xylose and fucose residues as well as those of S_4-RNase.

In the short type glycans were attached to Asn 18 of S_3- and S_5-RNases. All the *N*-glycans in the PS1 region including the HV region (Figure IV-2; Chapter V) were the short type glycans. Asn 71 of S_2-RNase had complex type with xylose residue and high mannose type sugar chains. Asn 74 of S_1- and S_4-RNases had complex type sugar chain with xylose and fucose residues. These residues were located at edge of the PS2 region (Chapter V). The only conserved *N*-glycosylation site among rosaceous S-RNase, Asn 121 in the Figure IV-2, had the complex, high mannose, and hybrid type sugar chains with or without xylose residue. In the PS3 region, all the sites except Asn 133 of S_2-RNase were glycosylated by the short type sugar chains. Asn 133 of S_2-RNase had the complex, high mannose, and hybrid type sugar chains with or without xylose residue. Asn 175 of S_2-RNase and Asn 176 of S_6-RNase near the *C*-terminus had the short type sugar chains. But Asn 175 of S_2-RNase had the complex type sugar chain with xylose and fucose residues as a major component.

Discussion

Analysis of sugar chains prepared from *P. pyrifolia* S-RNases revealed that *N*-glycans of these proteins were highly diversified. The complex type glycan with (or without) a β1->2 linked xylose and a α1->3 linked fucose residues, the high mannose and hybrid type glycans with (or without) xylose residue were observed in S-RNases. In addition, the short type glycan such as an *N*-acetylglucosamine and a chitobiose were detected. The complex type glycans

with xylose and fucose residues and xylomannose type glycans have been observed in glycoproteins from plant tissue, and the high mannose type glycan from plant and animal. The hybrid type glycans with xylose residue, which are minor component in S-RNase, are rarely observed. An *N*-acetylglucosamine is a rare structure, and a chitobiose as *N*-glycan has not been reported so far. These short type glycans were not produced by degradation of sugar chains during purification of S_4-RNase or manipulation of reductive pyridylamination of sugar chains, because each purification step was performed under cold temperature (4 °C) and the molecular weight of the purified S_4-RNase estimated by SDS-PAGE was equal to that estimated by 2D-PAGE in which S_4-RNase was directly extracted with lysis buffer (O'Farrell, 1975; Ishimizu *et al.*, 1996a).

Intermediate sugar chains between the complex or high mannose type sugar chains and the short type sugar chains such as M1A, M2B, M1X, and M2X were not detected in this study. Considering these results and biosynthetic pathway of *N*-glycan of glycoprotein, the short type glycans in S-RNases were likely to be produced by the action of endo-β-mannosidase to high mannose and complex type *N*-glycans and *N*-acetyl-β-glucosaminidase to a chitobiose. Endo-β-mannosidase, which has not ever found in any tissues, appears to exist in the style of *P. pyrifolia*.

The types of sugar chains depended on the position of glycosylation site. The conserved *N*-glycosylation site, Asn 121 in Figure IV-2, had the complex and high mannose type sugar chains with or without xylose residue. Since the remarkable difference of the structures and proportion of glycan at this site among seven S-RNases were not observed, the glycans of this site do not appear to serve as the discrimination of *S*-allele of pollen. In the tertiary structure of rosaceous S-RNase modeled based on that of fungus RNase Rh (Chapter V, Kurihara *et al.*, 1996), this asparagine residue is located at α-helix forming core structure of RNase, suggesting that the glycan of this site may play an important

role in the folding of core structure of these RNases. The function of this glycan may be similar to that of a sole conserved *N*-glycan near the *N*-terminus of solanaceous S-RNases. This glycan of solanaceous S-RNases is not necessary to self-incompatibility interaction (Karunanandaa *et al.*, 1994), and its function is still unknown.

Complex type sugar chains with xylose and fucose residues were observed only at Asn 74 of S_1-RNase, Asn175 of S_2-RNase, Asn 74 of S_4-RNase, and Asn 99 of S_7-RNase. This structure was not detected in *N. alata* S-RNases (Oxley and Bacic, 1995; Oxley *et al.*, 1996). Except for the sites of S_2-RNase, the other sites than the sites described above had only the short type sugar chains. Asn 18 of S_5-RNase and Asn 133 of S_4-RNase were not partially glycosylated. These structural difference of sugar chains among the sites appear to be due to the difference of the local environment and accessibility of glycosyltransferase and glycosidase to each glycosylation site. The processing enzyme to convert complex of high mannose type sugar chains to the short type sugar chain appeared to easily access to the sugar chains in the PS1 region and not to the sugar chain at the conserved site, Asn 121. The relationship between the type of sugar chain at each site and accessibility of processing enzyme to sugar chains was investigated (Faye *et al.*, 1986).

In S_2-RNase, the short type sugar chain was only 2 % of the whole, but Asn 133 of S_4-RNase and Asn 176 of S_6-RNase, of which position are respectively same as Asn 133 and Asn 175 of S_2-RNase, had only the short type sugar chains. Probably, it is difficult for the processing enzyme to access to the glycan of S_2-RNase. Ribonuclease activity against tolura yeast RNA of S_2-RNase was 10 to 20 times weaker than those of the other six S-RNases (unpublished result). The complex and high mannose type sugar chains of S_2-RNase may influence on the binding of substrate RNA to the active site of this protein as in the case of RNase B (Rudd *et al.*, 1994; Woods *et al.*, 1994).

Homology of the amino acid sequences between S_3- and S_5-RNases is 95.5 % and nine amino acid substitutions between these two S-RNases are located at the restricted region ranging from 21st to 90th (amino acid numbers) including the HV region. The *N*-glycosylation sites of these S-RNases are the same (Asn 18 and Asn 116). The structures and their proportion at each site of these two S-RNases are very similar (Tables 5 (c) and (d)). These results indicate that not carbohydrate moiety but amino acid moiety in the restricted region of these two S-RNase serve as discrimination between S_3 and S_5 pollen.

S-RNase is considered to be responsible for control of *S*-allele in the self-incompatibility reaction and be directly involved in the discrimination of *S*-allele. Since the location of *N*-glycosylation sites were different from every S-RNase, the carbohydrate moiety of S-RNase appears to serve as discrimination of *S*-allele of pollen. But since it is necessary that more than a threshold level of S-RNase expresses in the pistil to function as an factor of self-incompatibility (Lee *et al.*, 1994), heterogeneous carbohydrate moiety as a whole of S-RNase does not appear to be involved in the discrimination of *S*-allele, supposing that discrimination reaction by S-RNase was carried out via a molecule which recognize a specific structure.

From analogy of antigen binding site of MHC proteins, it was deduced that amino acid moieties of PS1, PS2, PS3, and PS4 regions, in which nonsynonymous substitution rate (d_N) was higher than synonymous substitution rate (d_S), were located at molecular surface and involved in the interaction of other molecule (Chapter V). But many sugar chains were attached to these region, especially PS1 and PS3 region. It is possible that *N*-glycans in these regions come into contact with the molecule interacted with S-RNase (if any). *N*-glycans at almost all the site are heterogeneous but it is considered that the common structure at each glycosylation site, such as reducing-end *N*-acetylglucosamine, is involved in the discrimination of *S*-alleles of pollen.

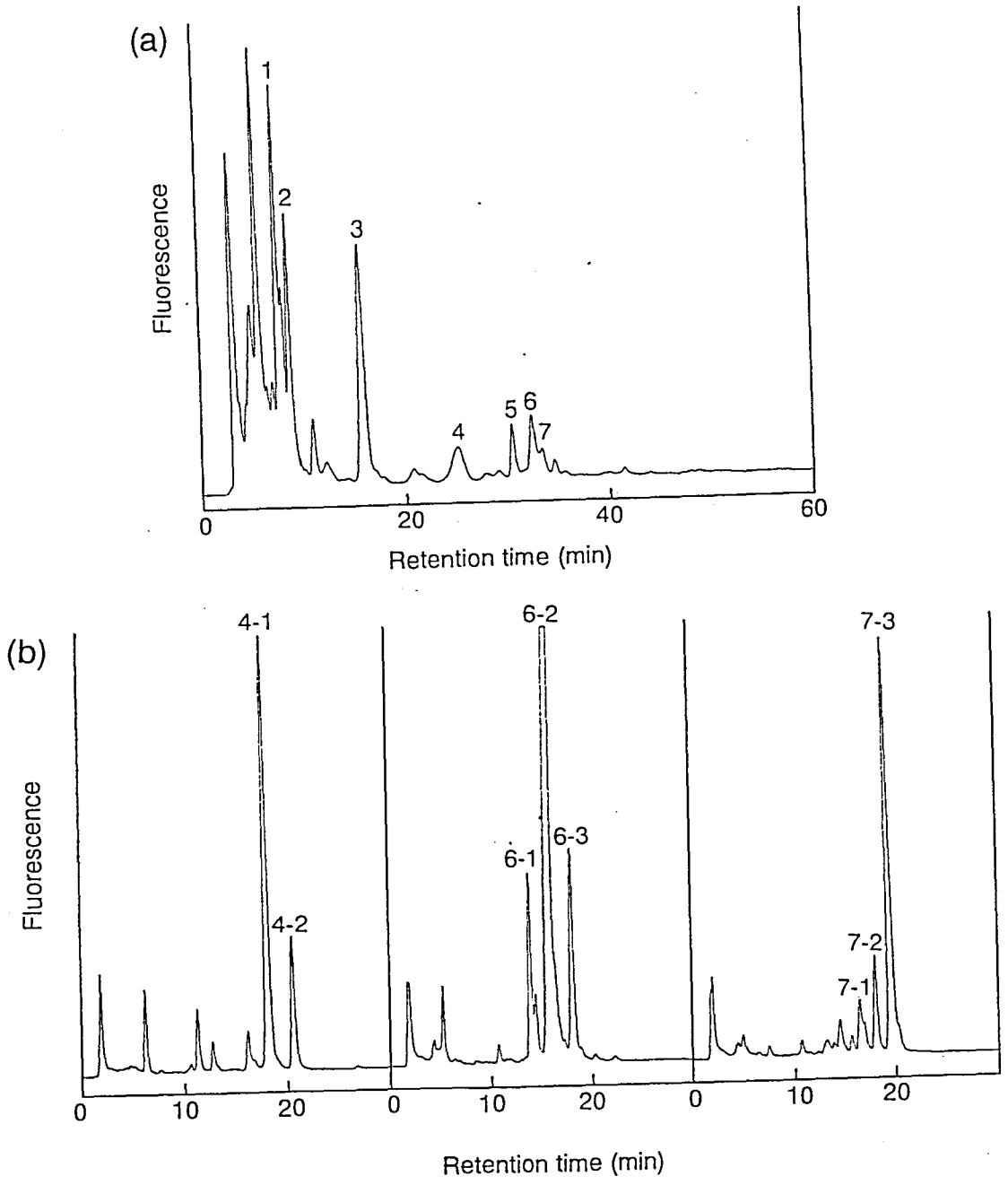


Figure IV-1. **Separation by (a) reversed-phase and (b) size-fractionation HPLCs of PA-sugar chains from S_4-RNase.** (a) The PA-sugar chains from S_4-RNase were separated by reversed-phase HPLC under the condition as described in the text. The fractions marked with numbers contained PA-sugar chains. (b) The fractions (4, 6, and 7) contained more than two PA-sugar chains in (a) were rechromatographed by size-fractionation HPLC under the condition as described in the text.

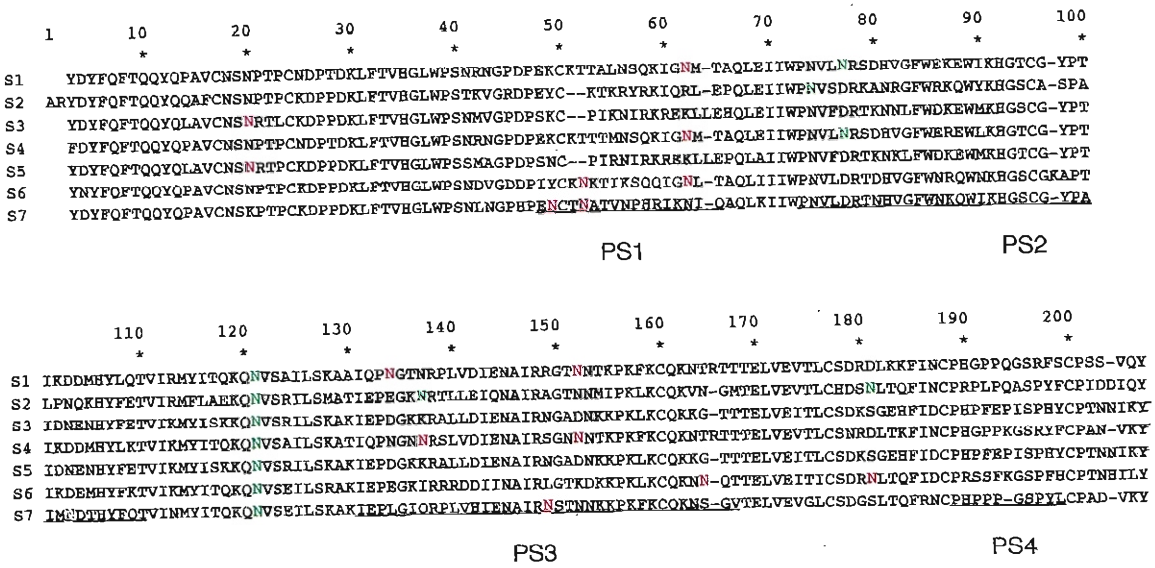


Figure IV-2. **Amino acid sequences and *N*-glycosylation sites of *P. pyrifolia* S-RNases.** Sequences were aligned manually. *N*-glycosylation sites are colored. The asparagine residues having only the short type sugar chains are shown in red. The asparagine residues bearing the complex or high mannose type sugar chains are shown in green.

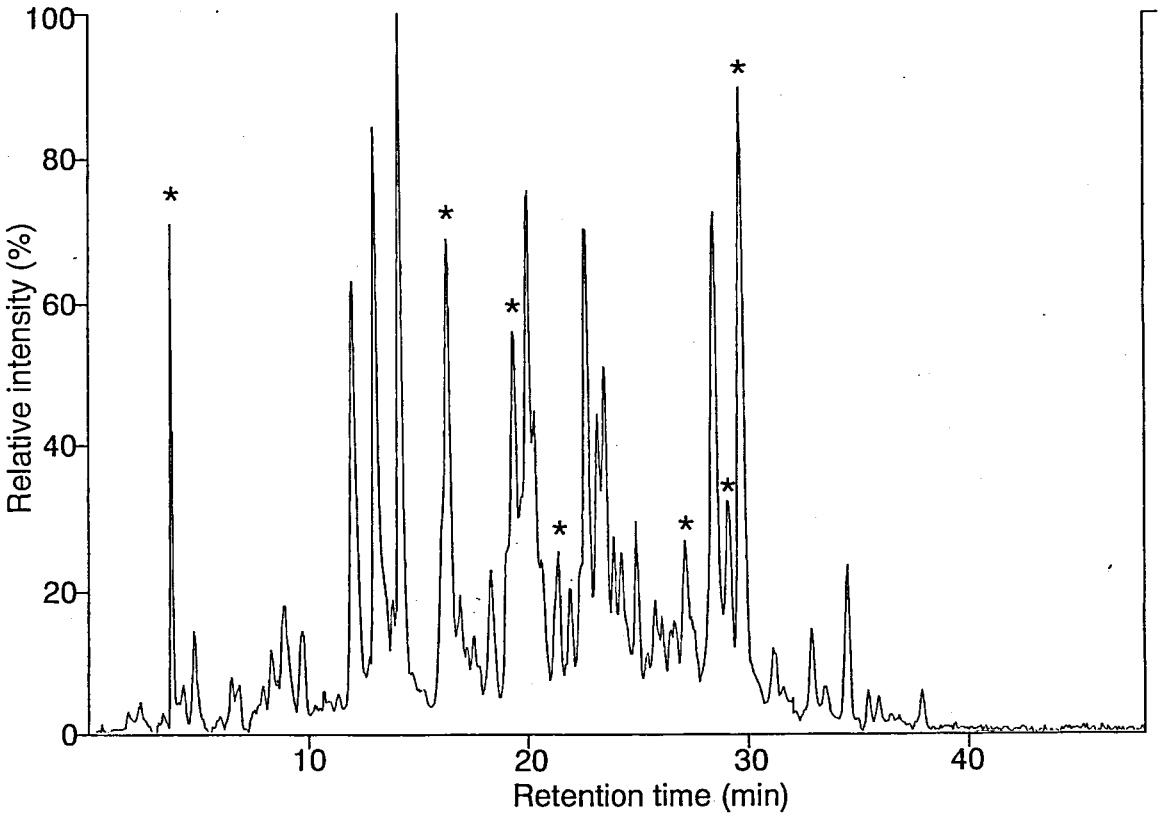
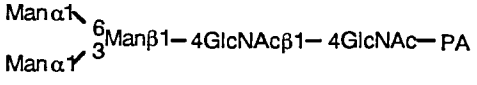
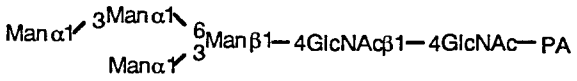
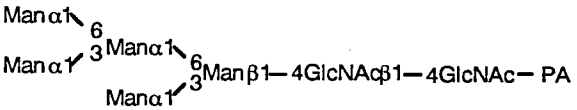
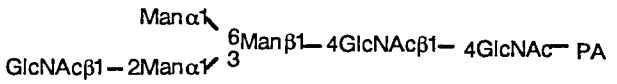
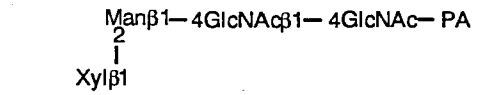
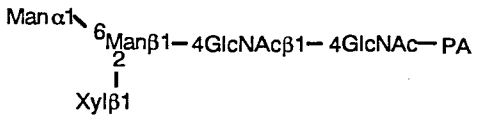
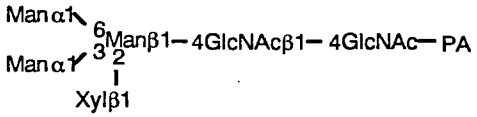
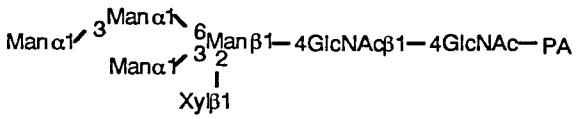
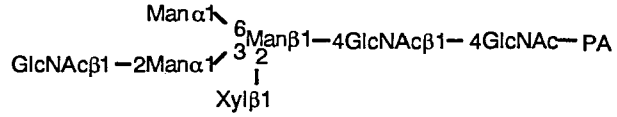
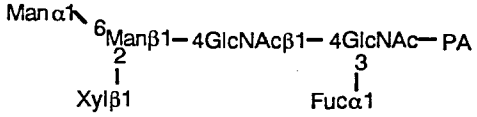
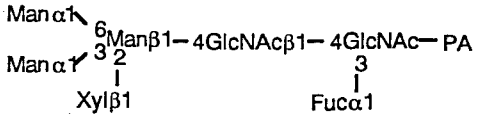
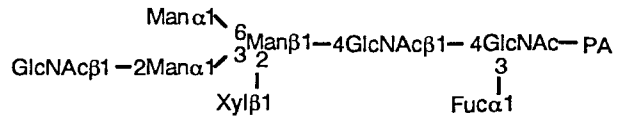


Figure IV-3. **LC / ESI-MS-recorded chromatogram for the protease digest of RCM-S_4-RNase.** RCM-S_4-RNase was digested with V8 and API, then chromatographed on a PLRP-S column (1.0 mm x 150 mm). Peptides were monitored by total ion current. Scans from 300 Da to 2400 Da were made every 4.2 sec. The glycopeptides were marked with asterisks.

Table IV-1. **Structures, abbreviations, reversed-phase scale (RPS), and molecular size (glucose unit) of authentic PA-sugar chains used in this study.**

Structure	Abbreviation	RPS	GU
ManNAc—PA	MN	12.9	0.8
GlcNAc—PA	GN	15.5	0.9
GlcNAcβ1—4GlcNAc—PA	GN2	27.4	1.7
Manβ1—4GlcNAcβ1—4GlcNAc—PA	M1	34.0	2.5
	M2B	42.6	3.3
	M3B	43.1	4.4
	M4C	44.9	5.3
	M5A	45.6	6.4
	AG1	41.1	5.3
	MX	41.1	2.8
	M2X	47.2	3.7
	M3X	42.6	5.1
	M4X	44.3	5.8
	AG1X	40.2[a]	5.8[b]
	M2FX	36.3	4.5
	M3FX	31.0	6.0
	AG1FX	30.6[a]	6.7[b]

[a]RPS for AG1X and AG1FX were calculated by the method of Yanagida et al. (1998).

[b]Molecular size for AG1X and AG1FX were calculated by the method of Hase et al. (1986).

Table IV-2. **RPS, molecular size, mono- and di- saccharide analysis by sugar AX-I column, and reducing end of PA-sugar chains isolated from S_4-RNase and their hydrolyzates.**

Fraction No.	State	RPS[a]	molecular size[b]	Sugar AX-I	Proposed structure[c]	Redusing-end[d]
1	a. intact	15.4	1.0	GN	GN	GN
2	a. intact	20.6	1.8	(GN-MN)	GN-MN	MN
	b. NAcG'ase[e] digestion of a	13.0	0.8	MN	MN	
3	a. intact	27.4	1.7	GN2	GN2	GN
	b. NAcG'ase digestion of a	15.7	0.9	GN	GN	
4-1	a. intact	30.8	5.9		M3FX	GN
	b. Partial acid hydrolysis of a	43.7	4.4		M3B	
	c. Man'ase[f] digestion of a	31.2	3.5		MFX	
	d. Fuc'ase[g] digestion of c	41.3	2.8		MX	
	e. Partial acid hydrolysis of d	34.1	2.5		M	
	f. Partial acid hydrolysis of d	27.5	1.6		GN2	
	g. Partial acid hydrolysis of d	15.3	0.9		GN	
4-2	a. intact	30.5	6.7		AG1FX	GN
	b. Partial acid hydrolysis of a	41.0	5.3		AG1	
	c. NAcG'ase digestion of a	31.1	6.0		M3FX	
	d. Man'ase digestion of c	30.6	3.5		MFX	
	e. Fuc'ase digestion of d	40.6	2.8		MX	
	f. Partial acid hydrolysis of e	34.0	2.5		M	
	g. Partial acid hydrolysis of e	27.4	1.7		GN2	
	h. Partial acid hydrolysis of e	15.5	0.9		GN	
5	a. intact	36.1	4.4		M2FX	GM
	b. Partial acid hydrolysis of a	42.8	3.3		M2B	
	c. Man'ase digestion of a	30.6	3.5		MFX	
	d. Fuc'ase digestion of c	40.7	2.8		MX	
	e. Partial acid hydrolysis of d	33.8	2.5		M	
	f. Partial acid hydrolysis of d	27.4	1.6		GN2	
	g. Partial acid hydrolysis of d	15.5	0.9		GN	
6-1	a. intact	43.8	4.4		M3B	GN
	b. Man'ase digestion of a	33.8	2.5		M	
	c. Partial acid hydrolysis of b	27.4	1.6		GN2	
	d. Partial acid hydrolysis of b	15.5	0.9		GN	

Fraction No.	State	RPS[a]	molecular size[b]	Sugar AX-I	Proposed structure[c]	Redusing-end[d]
6-2	a. intact	42.5	5.0		M3X	GN
	b. Partial acid hydrolysis of a	43.1	4.4		M3	
	c. Man'ase digestion of a	40.6	2.8		MX	
	d. Partial acid hydrolysis of c	33.6	2.5		M	
	e. Partial acid hydrolysis of c	27.0	1.6		GN2	
	f. Partial acid hydrolysis of c	15.3	0.9		GN	
6-3	a. intact	41.5	5.7		AG1X	GN
	b. Partial acid hydrolysis of a	41.5	5.3		AG1	
	c. NAcG'ase digestion of a	42.6	5.0		M3X	
	d. Man'ase digestion of c	40.5	2.8		MX	
	e. Partial acid hydrolysis of d	33.9	2.4		M	
	f. Partial acid hydrolysis of d	27.3	1.7		GN2	
	g. Partial acid hydrolysis of d	15.2	0.9		GN	
7-1	a. intact	44.9	5.2		M4C	GN
	b. Man'ase digestion of a	33.5	2.5		M	
	c. Partial acid hydrolysis of b	27.3	1.7		GN2	
	d. Partial acid hydrolysis of b	15.4	0.9		GN	
7-2	a. intact	43.8	5.7		M4X	GN
	b. Partial acid hydrolysis of a	44.8	5.3		M4C	
	c. Man'ase digestion of a	40.5	2.8		MX	
	d. Partial acid hydrolysis of c	34.1	2.4		M	
	e. Partial acid hydrolysis of c	27.4	1.6		GN2	
	f. Partial acid hydrolysis of c	15.3	0.9		GN	
7-3	a. intact	45.0	6.3		M5A	GN
	b. Man'ase digestion of a	33.6	2.5		M	
	c. Partial acid hydrolysis of b	27.2	1.6		GN2	
	d. Partial acid hydrolysis of b	15.4	0.9		GN	

[a]Reversed-phase scales (RPS) calculated by the method of Yanagida et al. (1998).

[b]Molecular size (glucose units) estimated by size-fractionation HPLC.

[c]Structures of PA-sugar chains proposed by 2D sugar map profile listed in Table I.

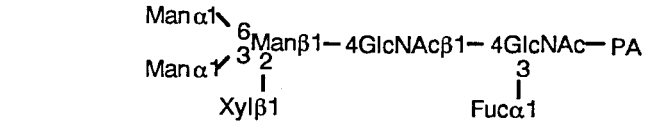
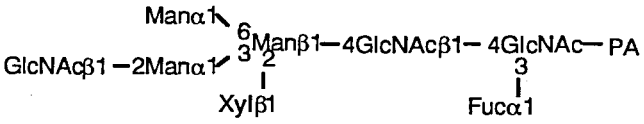
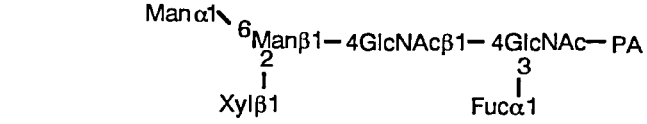
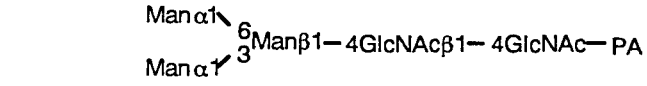
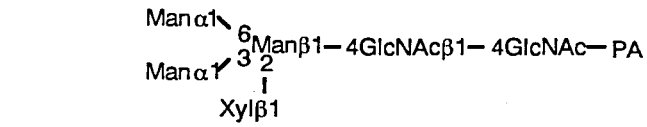
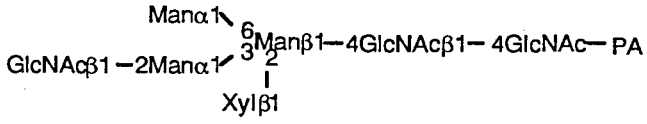
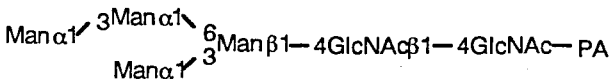
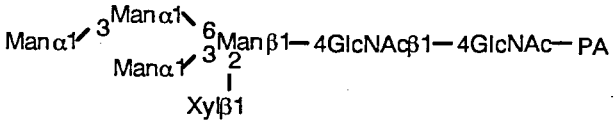
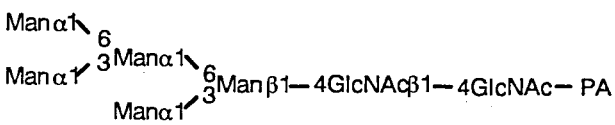
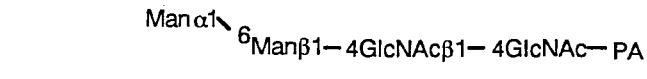
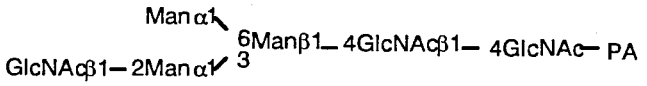
[d]Structures of reducing-end sugar identified as corresponding PA-sugar chains.

[e]N-Acetyl-β-D-glucosaminidase (*Diplococcus pneumoniae*)

[f]α-Mannosidase (Jack bean)

[g]α-L-Fucosidase (*Charonia lampus*)

Table IV-3. **Proposed structures for sugar chains of S_4-RNase and their relative amount.**

Fraction No.	Structure	Abbrevi -ation	Ratio (%) (by HPLC)[a]	Ratio (%) (by LC / MS)[b]
1	GlcNAc– PA	GN	15.9	17.8
2	GlcNAcβ1– ManNAc– PA	GN-MN	7.8	} 55.5
3	GlcNAcβ1– 4GlcNAc– PA	GN2	54.0	
4-1		M3FX	7.0	2.3
4-2		AG1FX	2.2	0.9
5		M2FX	5.5	0.7
6-1		M3B	0.9	2.1
6-2		M3X	7.5	10.7
6-3		AG1X	1.4	0.9
7-1		M4C	0.4	1.6
7-2		M4X	0.5	1.8
7-3		M5A	2.5	4.8
			n.d.	0.5
			n.d.	0.7

[a, b]Relative amount of sugar chains were calculated from the peak area observed by reversed-phase HPLC or from the ion intensity observed by LC / MS.

Table IV-4. **Sugar compositions of the *N*-linked glycopeptides from RCM-S_4-RNase.**

Site	Peptide position	Retention time	Molecular mass		Expected sugar composition	ratio
			calculated	observed		
		min	Da			
Asn60	58-66	19.4	1179.3	1179.1	$HexNAc_1$	53
			1382.5	1382.0	$HexNAc_2$	47
Asn74	67-85	27.2	2774.0	2774.2	$HexNAc_2$	(100)
	67-83	29.7	2285.6	2285.0	$HexNAc_1$	4
			2488.7	2488.4	$HexNAc_2$	79
		29.2	3091.3	3091.6	$HexNAc_2Hex_2deoxyHex_1Pen_1$	3
			3253.4	3253.2	$HexNAc_2Hex_3deoxyHex_1Pen_1$	10
			3456.6	3457.3	$HexNAc_3Hex_3deoxyHex_1Pen_1$	4
Asn117	116-124	16.0	1689.8	1689.2	$HexNAc_2Hex_2$	2
			1851.9	1851.2	$HexNAc_2Hex_3$	10
			1984.0	1984.0	$HexNAc_2Hex_3Pen_1$	51
			2014.1	2014.0	$HexNAc_2Hex_4$	7
			2055.1	2054.4	$HexNAc_3Hex_3$	3
			2176.2	2175.6	$HexNAc_2Hex_5$	22
			2187.3	2186.8	$HexNAc_3Hex_3Pen_1$	5
Asn133	125-140	21.3	1740.9	1741.7	none	73
			1994.1	1944.8	$HexNAc_1$	4
			2147.3	2150.0	$HexNAc_2$	22
Asn148	141-151	3.8	1594.7	1594.5	$HexNAc_2$	100

Table IV-5. **Assignment of molecular masses to each (glyco)peptide of API digests of S-RNases.** (a) S_1-RNase, (b) S_2-RNase, (c) S_3-RNase, (d) S_5-RNase, (e) S_6-RNase, (f) S_7-RNase.

(a)

Site	Peptide Position	Retention time	Molecular mass		Expected sugar composition	Ratio
			calculated	observed		
		min	Da			%
Asn 60 Asn 74	58-84	28.2	3779.2	3778.5	$HexNAc_3$	20
			3982.4	3981.3	$HexNAc_4$	9
		27.8	4584.0	4582.8	$HexNAc_4Hex_2deoxyHex_1Pen_1$	28
			4746.1	4745.8	$HexNAc_4Hex_3deoxyHex_1Pen_1$	43
Asn 117	116-124	14.7	1851.9	1851.2	$HexNAc_2Hex_3$	21
			1984.0	1984.4	$HexNAc_2Hex_3Pen_1$	53
			2014.0	2014.4	$HexNAc_2Hex_4$	2
			2055.1	2055.2	$HexNAc_3Hex_3$	7
			2146.2	2145.6	$HexNAc_2Hex_4Pen_1$	3
			2176.2	2176.0	$HexNAc_2Hex_5$	6
			2187.2	2186.4	$HexNAc_3Hex_3Pen_1$	8
Asn 130 Asn 148	125-151	21.1	3543.9	3544.1	$HexNAc_3$	11
			3747.1	3748.4	$HexNAc_4$	89

(b)

Site	Peptide position	Retention time	Molecular mass		Expected sugar composition	Ratio
			calculated	observed		
		min.	Da			%
Asn 71	58-76	22.8	3197.5	3197.5	$HexNAc_2Hex_2Pen_1$	5
			3359.6	3359.8	$HexNAc_2Hex_3Pen_1$	15
			3551.8	3551.4	$HexNAc_2Hex_5$	3
			3562.8	3562.6	$HexNAc_3Hex_3Pen_1$	77
Asn 117	116-132	20.8	2766.0	2765.7	$HexNAc_2Hex_3$	5
			2898.1	2898.0	$HexNAc_2Hex_3Pen_1$	36
			2928.1	2929.2	$HexNAc_2Hex_4$	2
			2969.2	2968.5	$HexNAc_3Hex_3$	5
			3090.3	3090.3	$HexNAc_2Hex_5$	5
			3101.3	3100.8	$HexNAc_3Hex_3Pen_1$	38
			3252.4	3252.3	$HexNAc_2Hex_6$	8
Asn 133	133-153	22.4	3085.4	3082.0	$HexNAc_2Hex_2$	4
			3247.6	3245.8	$HexNAc_2Hex_3$	11
			3379.7	3377.3	$HexNAc_2Hex_3Pen_1$	14
			3541.8	3540.0	$HexNAc_2Hex_4Pen_1$	60
			3745.0	3741.9	$HexNAc_3Hex_4Pen_1$	11
Asn 175	159-201	33.8	5464.1	5464.5	$HexNAc_2$	9
		33.2	6066.6	6066.7	$HexNAc_2Hex_2deoxyHex_1Pen_1$	5
			6228.8	6229.9	$HexNAc_2Hex_3deoxyHex_1Pen_1$	42
			6432.0	6432.7	$HexNAc_3Hex_3deoxyHex_1Pen_1$	45

(c)

Site	Peptide position	Retention time	Molecular mass		Expected sugar composition	Ratio
			calculated	observed		
		min	Da			%
Asn 18	1-23	25.4	3153.5	3153.0	$HexNAc_1$	37
			3356.7	3356.7	$HexNAc_2$	63
	114-123	8.5	1903.1	1902.6	$HexNAc_2Hex_2$	4
			2035.2	2035.6	$HexNAc_2Hex_2Pen_1$	6
			2065.2	2064.9	$HexNAc_2Hex_3$	9
			2197.3	2197.7	$HexNAc_2Hex_3Pen_1$	25
			2227.3	2226.0	$HexNAc_2Hex_4$	7
			2268.4	2268.7	$HexNAc_3Hex_3$	11
			2389.5	2389.2	$HexNAc_2Hex_5$	16
			2400.5	2400.2	$HexNAc_3Hex_3Pen_1$	23
Asn 116	115-122	13.1	1808.9	1809.6	$HexNAc_2Hex_3$	6
			1941.0	1940.8	$HexNAc_2Hex_3Pen_1$	31
			1971.0	1970.0	$HexNAc_2Hex_4$	5
			2012.1	2011.2	$HexNAc_3Hex_3$	19
			2133.2	2132.0	$HexNAc_2Hex_5$	15
			2144.2	2143.6	$HexNAc_3Hex_3Pen_1$	25
	115-123	10.9	1774.9	1774.9	$HexNAc_2Hex_2$	3
			1907.0	1907.0	$HexNAc_2Hex_2Pen_1$	7
			1937.0	1936.9	$HexNAc_2Hex_3$	13
			2069.2	2069.2	$HexNAc_2Hex_3Pen_1$	23
			2099.2	2099.7	$HexNAc_2Hex_4$	4
			2140.2	2140.0	$HexNAc_3Hex_3$	13
			2261.3	2260.8	$HexNAc_2Hex_5$	13
			2272.3	2271.8	$HexNAc_3Hex_3Pen_1$	25

(d)

Site	Peptide position	Retention time	Molecular mass		Expected sugar composition	Ratio
			calulated	observed		
		min	Da			%
Asn 18	1-23	24.9	2934.2	2934.7	none	11
			3137.4	3137.7	$HexNAc_1$	36
			3340.6	3340.6	$HexNAc_2$	53
Asn 116	114-123	8.8	2035.2	2035.5	$HexNAc_2Hex_2Pen_1$	7
			2065.2	2064.9	$HexNAc_2Hex_3$	9
			2197.3	2197.4	$HexNAc_2Hex_3Pen_1$	37
			2268.4	2268.5	$HexNAc_3Hex_3$	11
			2389.5	2389.2	$HexNAc_2Hex_5$	7
			2400.5	2400.4	$HexNAc_3Hex_3Pen_1$	28
	115-122	13.3	1808.9	1809.2	$HexNAc_2Hex_3$	7
			1941.0	1940.4	$HexNAc_2Hex_3Pen_1$	27
			2012.1	2011.2	$HexNAc_3Hex_3$	20
			2133.2	2132.0	$HexNAc_2Hex_5$	18
			2145.2	2144.4	$HexNAc_3Hex_3Pen_1$	28
	115-123	11.1	1774.9	1774.6	$HexNAc_2Hex_2$	6
			1907.0	1906.9	$HexNAc_2Hex_2Pen_1$	6
			1937.0	1937.3	$HexNAc_2Hex_3$	12
			2069.2	2069.1	$HexNAc_2Hex_3Pen_1$	35
			2140.2	2140.6	$HexNAc_3Hex_3$	17
			2272.3	2272.0	$HexNAc_3Hex_3Pen_1$	25

(e)

Site	Position	Retention time	Molecular mass		Expected sugar composition	Ratio
			calculated	observed		
		min.	Da			%
Asn50	50-51	2.2	463.5	463.2	$HexNAc_1$	66
			666.7	666.4	$HexNAc_2$	34
Asn60	55-88	30.3	4265.8	4266.9	$HexNAc_1$	56
			4469.0	4469.1	$HexNAc_2$	44
Asn118	117-127	15.6	1975.1	1975.3	$HexNAc_2Hex_2$	3
			2137.2	2137.4	$HexNAc_2Hex_3$	10
			2269.4	2268.7	$HexNAc_2Hex_3Pen_1$	3
			2299.4	2299.6	$HexNAc_2Hex_4$	14
			2340.4	2340.5	$HexNAc_3Hex_3$	11
			2461.5	2461.1	$HexNAc_2Hex_5$	33
			2502.6	2503.1	$HexNAc_3Hex_4$	15
			2664.7	2664.5	$HexNAc_3Hex_5$	11
Asn161 Asn176	160-189	27.7	3996.4	3996.5	$HexNAc_2$	11
			4199.6	4199.3	$HexNAc_3$	55
			4402.8	4403.5	$HexNAc_4$	34

(f)

Site	Peptide position	Retention time	Molecular mass		Expected sugar composition	Ratio
			calculated	observed		
		min	Da			%
Asn47 Asn50	29-59	21.0	3929.3	3930.0	$HexNAc_2$	15
			4132.5	4134.1	$HexNAc_3$	52
			4335.7	4337.2	$HexNAc_4$	33
Asn99	89-115	29.3	4200.5	4202.0	$HexNAc_2Hex_2deoxyHex_1Pen_1$	15
			4362.7	4362.6	$HexNAc_2Hex_3deoxyHex_1Pen_1$	62
			4565.9	4565.6	$HexNAc_3Hex_3deoxyHex_1Pen_1$	23
Asn117	116-125	15.4	1910.0	1910.0	$HexNAc_2Hex_3$	25
			2042.1	2043.0	$HexNAc_2Hex_3Pen_1$	29
			2072.1	2072.0	$HexNAc_2Hex_4$	10
			2113.2	2113.0	$HexNAc_3Hex_3$	11
			2204.2	2205.0	$HexNAc_2Hex_4Pen_1$	8
			2245.3	2245.0	$HexNAc_3Hex_3Pen_1$	9
			2275.3	2275.0	$HexNAc_3Hex_4$	3
			2407.4	2407.0	$HexNAc_3Hex_4Pen_1$	5
Asn145	127-150	21.8	2930.3	2930.6	$HexNAc_1$	25
			3133.5	3134.0	$HexNAc_2$	75

Chapter V
Identification of the regions in which positive selection may operate in rosaceous S-RNases

Introduction

In this chapter, the primary structures of rosaceous S-RNases were further characterized from a view of molecular evolutionary genetics. Rare *S*-alleles have a reproductive advantage because pollen bearing such alleles is less likely to land on a stigma with the same allele, and many kinds of *S*-alleles are maintained in a finite population (Wright, 1939). Overdominant selection (heterozygote advantage) therefore is considered to occur at the *S*-locus in the population of the genus. Recently, recognition sites in some proteins have been reported to be regions in which the number of nonsynonymous nucleotide substitutions (d_N) exceeds that of synonymous substitutions (d_S), and that positive selection probably takes place in these regions. For example, the antigen recognition site of MHC proteins and the antigenic epitopes of the antigenic surface proteins of parasites and viruses are known to be regions with an excess of d_N over d_S (Hughes and Nei, 1988, 1989; Endo *et al.*, 1996). It has been considered positive selection may operated in these regions. Proteins related to SI, including S-RNase, will have such a region if overdominant selection operates in them (Maruyama and Nei, 1981).

In the Solanaceae, the primary structural features of S-RNases have been studied. Pairwise comparisons show that these S-RNases have highly divergent amino acid sequences and that some interspecific pairs have higher sequence

similarities than the intraspecific pairs (Ioerger *et al.*, 1990). The level of constraint on nucleotide substitution is heterogeneous throughout the S-RNase gene, some regions being highly constrained others virtually unconstrained (Clark and Kao, 1991). Window analysis of d_N and d_S, however, failed to detect a region in which d_N exceeds d_S. Moreover, no region with an excess of d_N over d_S was detected in S-related Brassicaceae proteins (Hinata *et al.*, 1995). No recognition sites have yet been identified even in series of transgenic experiments using chimeric genes between two S-RNases (Kao and McCubbin, 1996; Zurek *et al.*, 1997).

Several cDNAs encoding S-RNases from rosaceous species (*P. pyrifolia* and *M. domestica*) have been cloned (Broothaerts *et al.*, 1995; Janssens *et al.*, 1995; Norioka *et al.*, 1995; Sassa *et al.*, 1996; Sassa and Hirano, 1997; Chapter II, Ishimizu *et al.*, 1998b). These S-RNases are distinct from the solanaceous and scrophulariaceous S-RNases and other RNase T_2 family enzymes in sequence alignment (Broothaerts *et al.*, 1995; Norioka *et al.*, 1996; Sassa *et al.*, 1996) or in the neighbor-joining phylogenetic tree (Xue *et al.*, 1996). To identify the *S*-allele-specific recognition site in S-RNase, I conducted window analysis of d_S and d_N in 11 maloideous S-RNases and found four regions with an excess of d_N over d_S in them. Moreover, their secondary structures predicted by the PHD method were very similar to the RNase Rh structure, the only known tertiary structure ribonuclease in the RNase T_2 family. I here discusses *S*-allele-specific recognition sites in the rosaceous S-RNases on the basis of the above findings.

Materials and Methods

Sequences of S-RNases

The nucleotide and amino acid sequences of the maloideous S-RNases used were *P. pyrifolia* S_2- and S_4-RNases (D49527 and D49528) (Norioka *et al.*, 1995); S_1-, S_3-, S_5-, S_6-, and S_7-RNases (AB002139, AB002140, AB002141, AB002142, and AB002143) (Ishimizu *et al.*, 1998b); *M. domestica* S_2-, S_3-RNases (U12199, U12200) (Broothaerts *et al.*, 1995), S_7-, S_9-RNases (U19792, U19793) (Janssens *et al.*, 1995); and S_f-RNase (D50837) (Sassa *et al.*, 1996). Those of the solanaceous S-RNases used are *Lycopersicon peruvianum* S_3-RNases (X76065) (Royo *et al.*, 1994); *Nicotiana alata* S_6-RNase (U08861) (Anderson *et al.*, 1989); and *Antirrhinum hispanicum* S_2-RNases (X96465) (Xue *et al.*, 1996).

Prediction of secondary structures

Secondary structures of the maloideous, solanaceous, and scrophulariaceous S-RNases were predicted by the PHD method (Rost and Sander, 1993).

Sequence alignment and window analysis of d_S and d_N

On the basis of the predicted secondary structures, the amino acid sequences of the S-RNases were aligned manually with the sequence of *Rhizopus niveus* RNase Rh (D12476) (Horiuchi *et al.*, 1988). The d_S and d_N values for each window of 20 codons along the aligned sequences in the S-RNase pairs were estimated (Endo *et al.*, 1996; Nei and Gojobori, 1986). The averages of the values for each window for all 55 pairs of the 11 S-RNases were plotted against the location of the window. Site number 1 on the horizontal column in Figure V-2 shows the window from codon number 1 to 20.

Results

Prediction of the secondary structures of S-RNases

Secondary structures of 12 maloideous, two solanaceous, and one scrophulariaceous S-RNases were predicted by the PHD method (Rost and Sander, 1993) (Figure V-1). Positions of the predicted α-helical and β-stranded regions were similar. Many conserved residues appear to participate in the formation of these structures. The topology of the predicted secondary structures coincided with the structure of RNase Rh, the only known tertiary structure ribonuclease in the RNase T_2 family (Kurihara *et al.*, 1996). The amino acid sequences of the S-RNases were aligned with the sequence of RNase Rh using the predicted secondary structures (Figure V-1). The frameworks of the S-RNases also are likely to be very similar to the framework of RNase Rh because the secondary structure forms its core structure.

Window analysis of d_S and d_N

The averages of d_S and d_N for each window of 20 codons for all 55 pairs of the seven *P. pyrifolia* and four *M. domestica* S-RNases are plotted against the location of the window (Figure V-2). Sequence alignment is the same as in Figure V-1. Four statistically significant regions in which d_N was higher than d_S were detected in site (amino acid) numbers 38-55 (48-65), 63-100 (73-110), 121-158 (131-168), and 179-188 (189-198). These regions respectively were designated PS1, PS2, PS3, and PS4 regions (regions which positive selection may operate). Remarkably higher d_N than in the other regions was found in the PS1 region. This region includes the HV region (amino acid numbers 51-66) which has many amino acid substitutions in maloideous S-RNases and is thought to be one of the *S*-allele-specific recognition sites (Ishimizu *et al.*, 1998b). The HVa and HVb regions in solanaceous (Ioerger *et al.*, 1991)

respectively correspond to the PS1 and PS2 regions described above (Figure V-1). The four PS regions are shaded in Figure V-1.

Location of the PS regions in S-RNase

On the assumption that the tertiary structure of S-RNase is similar to that of RNase Rh, the PS1, PS2, PS3, and PS4 regions correspond to two surface sites on RNase Rh (Figure V-3 (a)) based on the sequence alignment in Figure V-1. The PS1 and PS2 regions form one site on the left side of the tertiary structure of RNase Rh in Figure V-3 (a). The PS3 and PS4 regions form the other site on the right side. The substrate RNA-binding cleft (Nakamura *et al.*, 1993) is located between these two sites.

Positions of the amino acid substitutions in two highly homologous pairs of S-RNases are marked on the tertiary structure of RNase Rh (Figure V-3(b) and 3(c)). All nine amino acid substitutions between the *P. pyrifolia* S_3- and S_5-RNases pairs (96% sequence identity) are located in the PS1 and PS2 regions or their adjacent substrate-binding cleft (Figure V-3 (b)). In the pair of *P. pyrifolia* S_1- and S_4-RNases (90% sequence identity), 11 of the 20 substitutions are in or around the PS3 and PS4 regions (Figure V-3 (c)), six being located in or around the PS1 and PS2 regions.

Discussion

S-RNase is associated with GSI as a stylar component encoded by the *S*-locus and proposed to recognize *S*-alleles. To determine where the *S*-allele-specific recognition sites in S-RNase are and how S-RNase acts in the self-incompatible reaction, I analyzed d_S and d_N for windows of 20 codons each of

pairs of 11 maloideous S-RNases. Four regions in which d_N exceeded d_S were detected in maloideous S-RNases (Figure V-2). This is the first report of regions with an excess of d_N over d_S in SI related proteins. No such region was detected in solanaceous S-RNases (Clark and Kao, 1991). This difference may be due to the ages of the respective gene groups because the d_S values of solanaceous S-RNases are about two fold those of maloideous S-RNases (Clark and Kao, 1991). In solanaceous S-RNases, however, the d_N values of regions corresponding to the four PS regions described above are slightly higher than those of the other regions (Clark and Kao, 1991). In addition, no region with an excess of d_N over d_S was detected in proteins associated with SSI, although some regions had high d_N values (Hinata *et al.*, 1995).

The d_N value in the PS1 region was considerably higher than those in the other regions, even in the other three PS regions. Insertion, deletion, or recombination of genes as well as nucleotide substitution may easily occur in this region. Interestingly, the only intron in the S-RNases is inserted between the 59th and 60th amino acid residues in this region (unpublished results). The difference in d_N between the PS1 and the other PS regions seems to depend on these regions having different functions. A large number of nucleotide substitutions may be necessary in the PS1 region for *S*-allele-specific discrimination.

Regions with an excess of d_N over d_S are present in the antigen recognition sites of MHC proteins (Hughes and Nei, 1988; Hughes and Nei, 1989; Endo *et al.*, 1996) and antigenic surface proteins of parasites and viruses (Endo *et al.*, 1996). Recently, McCubbin *et al.* (1997) reported an experimental evidence that S-RNase is a recognition molecule that interacts with its counterpart of pollen. The four regions in S-RNase also probably function as recognition sites that interact with certain pollen molecules, and positive selection may take place in these regions.

The tertiary structures of maloideous S-RNases are probably similar to that of RNase Rh judging from the similarity between the secondary structure of RNase Rh and the predicted secondary structures of the S-RNases. The tertiary structures of the maloideous S-RNases were also predicted by 3D-1D compatibility method using the program COMPASS (Matsuo *et al.*, 1996; Ota and Nishikawa, 1997), showing that they are highly compatible with the RNase Rh structure (Nishikawa, K. and Ota, M., personal communication). On the tertiary structure of RNase Rh, the PS1 region forms the surface site together with the PS2 region (Figure V-3 (a)). The PS3 and PS4 regions form the surface site just opposite to the PS1 and PS2 regions. The question arises whether S-RNase interacts with one molecule at the two sites or with two molecules at each site. The molecules that interact with S-RNase may be proteins encoded at the *S*-locus that are expressed specifically in pollen or products of modifier genes other than the *S*-locus (Uyenoyama, 1991; Ai *et al.*, 1991), but no such products have been identified as yet.

The PS regions have many polar, especially basic, amino acid residues. Arginine and lysine residues tend to be located on the surface of the protein (Miller *et al.*, 1987), indicative that the PS regions also are located on the surface of S-RNase. These basic amino acid residues may function in the interaction with a counterpart molecule.

The sequence identity between the *P. pyrifolia* S_3- and S_5-RNases is very high (96%), and only nine amino acid substitutions are located in or around the site formed by the PS1 and PS2 regions (Figure V-3(b)). This site therefore may be responsible for discriminating between S_3 and S_5 pollens and for triggering the self-incompatible reaction. The *P. pyrifolia* S_1- and S_4-RNases also are highly homologous (90% sequence identity). Although the 20 amino acid substitutions between this pair are dispersed throughout the primary structure, their positions are located in or around the two sites on the tertiary structure of

RNase Rh (Figure V-3 (c)). These findings support the hypothesis that the PS regions function as recognition sites for the discrimination of *S*-alleles and do not contradict the report that one segment of S-RNase does not determine *S*-allele specificity (Zurek *et al.*, 1997).

These findings do not conflict with the classical genetic theory of GSI (Wright, 1939). A newly functional *S*-allele introduced into a population is preferred to other *S*-alleles because pollen bearing the new *S*-allele is less likely to match one of the two *S*-alleles of the pistil. Loss by random genetic drift of a new *S*-allele occurs less often than loss of a neutral allele, and the population of the new *S*-allele is increased until equilibrium is reached. Such new *S*-alleles must be generated by nonsynonymous substitutions at *S*-allele-specific recognition sites on the *S*-locus. The PS regions, which have high d_N values, therefore are speculated to be *S*-allele-specific recognition sites. The functions of the PS regions must now be investigated by biochemical and biological methods to clarify the molecular mechanism of S-RNase-based GSI.

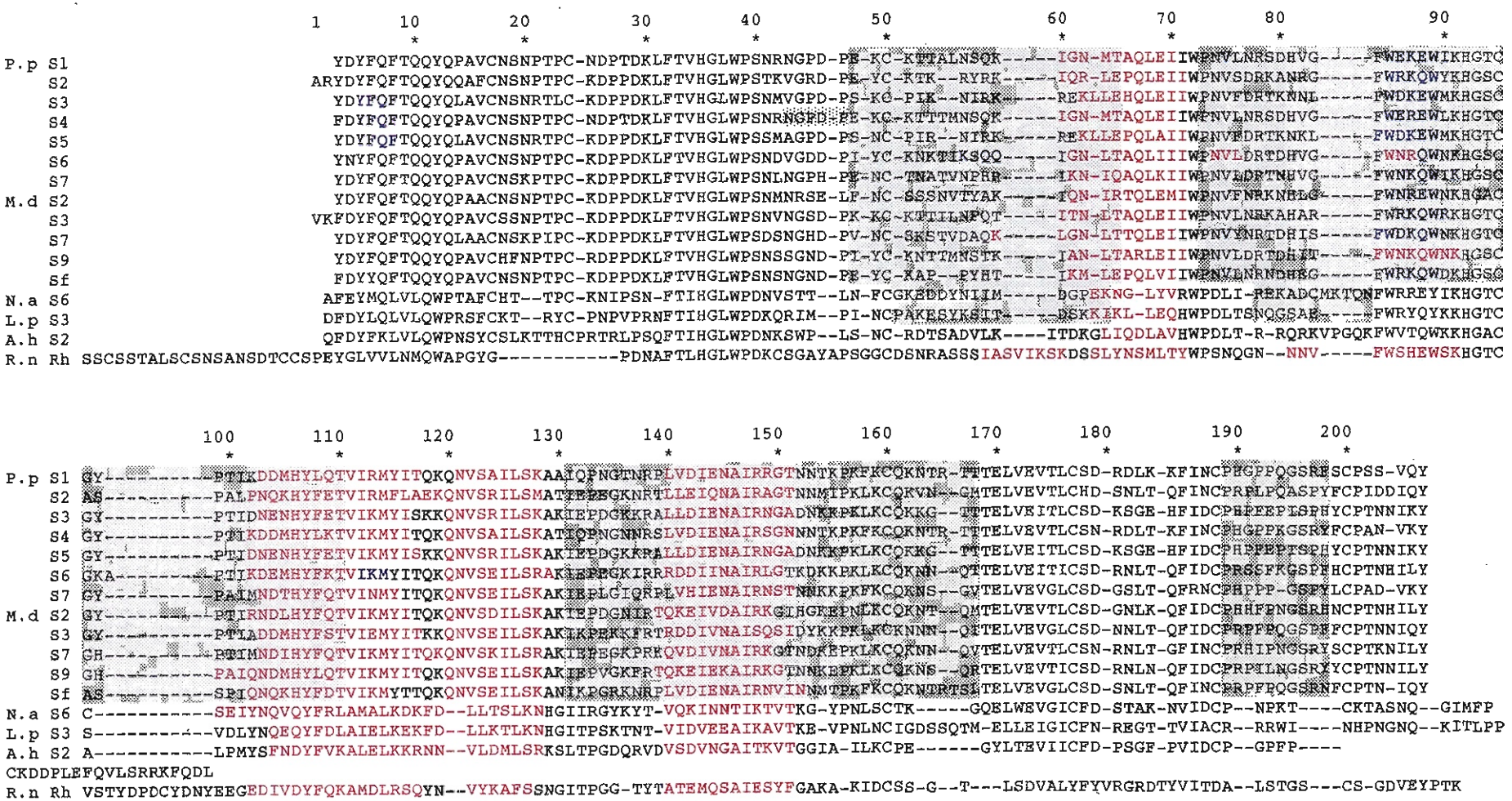


Figure V-1. **Predicted secondary structures of rosaceous S-RNases.** These structures were predicted by the PHD method (Rost and Sander, 1993), as were the structures of the solanaceous and scrophulariaceous S-RNases, *N. alata* S_6- (Anderson et al., 1989), *L. pervianum* S_3- (Royo et al., 1993), and *A. hispanicum* S_2-RNases (Xue et al., 1996). Residues predicted to form the α-helix are shown in red and those that form the β-strand in blue. The amino acid sequence of RNase Rh was aligned manually with the S-RNase sequences. The secondary structure of RNase Rh shown here was derived from its tertiary structure (Kurihara et al., 1996). Four PS regions in the rosaceous S-RNases are shaded, as are the HVa and HVb regions in the solanaceous S-RNases.

Figure V-2. **Window analysis of synonymous and nonsynonymous nucleotide substitutions in maloideous S-RNases.** The window is defined as a sequence region 20 codons long. Averages of the synonymous and nonsynonymous nucleotide substitutions for all 55 pairs of 7 *P. pyrifolia* and 4 *M. domestica* S-RNases were plotted against the location of the window. Synonymous substitutions are indicated by the blue line, nonsynonymous ones by the red one.

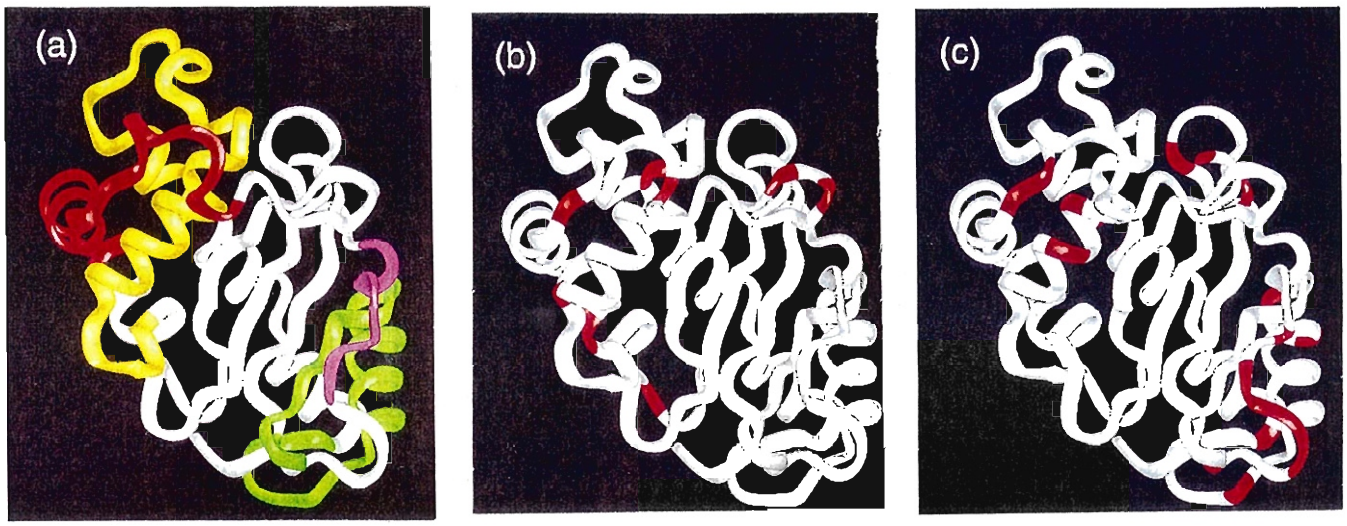


Figure V-3. **Locations of the PS regions of rosaceous S-RNases and the amino acid substitutions between the two highly homologous pairs of S-RNases on the tertiary structure of RNase Rh.** The PS1, PS2, PS3, and PS4 regions (a) respectively are colored red, yellow, green, and purple. Amino acid substitutions between (b) *P. pyrifolia* S_3- and S_5-RNases and (c) *P. pyrifolia* S_1- and S_4-RNases are shown in red.

General Discussion

Stylar S-RNase has been shown to be associated with the GSI of the Rosaceae (Broothaerts *et al.*, 1995; Sassa *et al.*, 1996; Norioka *et al.*, 1996; Chapter I, Ishimizu *et al.*, 1996a). S-RNase-based GSI also operates in the Solanaceae and Scrophulariaceae (Anderson *et al.*, 1986; Xue *et al.*, 1996). I studied the relationship between the structures and the function of the S-RNases. Each structure (the primary structure, the disulfide bond location, and the *N*-glycan structure) of the S-RNases was determined and characterized. The functions for discriminating between self and non-self pollen were discussed in separate chapters. What has not been discussed in detail in the previous chapters—the compatibility of the structural features of S-RNase obtained in this study with the proposed molecular mechanism of SI; the predicted characteristics of the unknown pollen S-products; the reasons for the high degree of polymorphism in the rosaceous S-RNases; the perspective and future direction of research on the molecular mechanism of the RNase-based GSI; and the application of the findings of this study to horticulture—are dealt with here.

How does S-RNase discriminate self and non-self pollen ?

Two models of *S*-allele-specific inhibition of pollen tube growth involving S-RNase have been proposed based on results of classical genetic and recent molecular biological experiments (Dodds *et al.*, 1996; Kao and McCubbin, 1996) (Figure 3). One is the *S*-allele-specific uptake model, the other the RNase inhibition model. At present there is not enough evidence to support either

model. The compatibility between the findings of my study and these two models is discussed.

S-RNase must have two functions, RNase activity and *S*-allele-specificity, to explain either model. In view of the tertiary structure of S-RNase, this stylar protein apparently consists of one conserved and two variable domains (Figure V-3 (a)). The one conserved domain, the uncolored domain in Figure V-3, has the substrate (RNA) binding cleft and the catalytic site of RNase activity (Ishimizu *et al.*, 1995). The two variable domains, the colored regions in Figure V-3 (a), located on both sides of this cleft, may be responsible for *S*-allele-specificity. These structural features of S-RNase do not conflict with the SI discrimination process of either model.

Genetic evidence supports the RNase inhibitor model (de Nettancourt, 1977). Assuming that the conserved S-RNase domain (the substrate RNA binding site) (Figure V-3) can interact with the pollen S-product nonspecifically, the S-RNase-based GSI is explained by the RNase inhibition model, i.e., the non-self dependent interaction model (Dodds *et al.*, 1996; Kao and McCubbin, 1996). In that model, pollen S-products interact with the conserved S-RNase domain nonspecifically and inhibit RNase activity; however, when the *S*-allele of the pollen S-product matches that of the S-RNase, it specifically binds to the variable S-RNase domain rather than to the conserved domain, resulting in the retention of RNase activity. As a result, the pollen S-product can inactivate all S-RNases except that of the same *S*-allele. Non-self-dependent dimerization has been reported for the two homeodomain proteins produced by the *b* locus in *Ustilago maydis* (Känmper *et al.*, 1995).

The discrimination of self and non-self pollen by S-RNase has been discussed on the assumption that a particular S-RNase region functions as the determinant for *S*-allele-specificity. The entire protein, however, may function in

S-allele-specificity because variable regions are spread all over the S-RNase structure (Chapters II and V).

Clearly, the molecules that interact with S-RNase have to be identified and characterized in order to assess the validity of the two models shown in Figure 3 or to propose alternative models.

Predicted characters of the molecule interacting with S-RNase

One candidate for the molecule(s) that interacts with the stylar S-RNase is a pollen S-protein which has yet to be found. S-RNase has *S*-allele-specific regions; therefore, pollen S-protein probably has an *S*-allele-specific domain, if the *S*-allele-specific uptake model is correct, or an RNase inhibitor domain and an *S*-allele-specific domain, if the RNase inhibitor model is correct. The pollen S-product ostensibly would be a ribonucleoprotein or RNA because the sides of the substrate binding cleft of S-RNase includes variable regions (Figure V-3). Moreover, because two variable domains are present in the model of the tertiary structure of the S-RNase (Figure V-3), the protein may interact with two molecules much the same as the MHC protein interacts with an antigen and the T cell receptor.

The determinant of the pollen *S*-phenotype is encoded by an *S*-locus gene other than the S-RNase gene (Sassa *et al.*, 1997). To date only a small region of the genomic DNA surrounding the S-RNase genes has been sequenced and characterized, but some pollen-expressed RNAs and proteins linked to the *S*-locus have been identified (Li *et al.*, 1997; Kao *et al.*, 1997; Kiyozumi, 1998). A gene called 48A that is specifically expressed in the pollen of *N. alata* (Li *et al.*, 1997) has been sequenced for some *S*-alleles, but a high degree of polymorphism has not been found in it. Its function in GSI has yet to be clarified.

What must be done to clarify the molecular mechanism of the RNase-based GSI ?

As for structural analysis of S-RNase, its tertiary structure should be determined in order to describe its function in detail. The functions of the S-RNase PS regions identified in this study must be analyzed by constructing transgenic plants that have an S-RNase gene mutated in the PS regions. If PS regions are responsible for interaction(s) with some molecule(s), they can be used to determine which molecule(s) interacts with S-RNase.

The S-RNase substrate, which seems to interact with the conserved domain of S-RNase, must be identified in order to clarify the molecular mechanism of SI. One candidate for this substrate is the pollen rRNA because the rRNA of the self pollen tube, but not that of the non-self pollen tube, is degraded in the pistil (McClure *et al.*, 1990). The RNase activities of S-RNases against yeast tRNA are weaker than those of the other RNase T_2 type ribonucleases (Takuma, 1996). This suggests that the actual in vivo substrate may be specific for S-RNase, as in the case of α-sarcin which nonspecifically cleaves RNA in vitro and cleaves rRNA at a specific site when the substrate is a ribosome (Endo *et al.*, 1983). Ribonucleases are associated with such varied phenomena as angiogenesis (Shapiro *et al.*, 1986), response to virus infection (Schneider *et al.*, 1993), and apoptosis (Castelli *et al.*, 1997). The substrates for these ribonucleases have yet to be identified, but they also may be similar to that for S-RNase.

How has a high degree of polymorphism been generated in the rosaceous S-RNase?

Rosaceous S-RNase is a highly polymorphic protein that has regions in which d_N is in excess of d_S (Chapters II and V, Ishimizu *et al.*, 1998b). This character is unique because the numbers of base substitutions in most eukaryotic genes of polymorphic alleles are very low and most substitutions are synonymous. This

character is expected to be expressed by a gene in which overdominant selection operates, such as genes that encode MHC proteins. In fact, MHC class I and II proteins have this character (Hughes and Nei, 1988, 1989). Overdominant selection also is considered to operate in SI-related proteins because rarely are homozygotes produced due to self-sterility (Wright, 1939).

How have polymorphic *S*-allelic ribonucleases been generated? The high degree of polymorphism is considered to be due to three factors: (1) a high mutation rate, (2) gene conversion or interlocus genetic exchange, and (3) overdominant selection (Hughes and Nei, 1988). The amino acid substitution sites in S_1- and S_4-RNases are found only in the C-terminal region and in S_3- and S_5-RNase in the *N*-terminal region (Figure II-4). Interlocus genetic exchange in the PS1 region may generate polymorphic proteins. To test this speculation, I also determined the nucleotide sequences of a sole intron inserted in the PS1 region in seven *P. pyrifolia* S-RNases (Ishimizu *et al.*, unpublished data) and calculated the number of base substitutions in each region (Chapter V; unpublished data). If a high mutation rate, gene conversion, or interlocus genetic exchange is responsible for S-RNase polymorphism, the number of base substitutions would be higher in a particular region than in other regions, and the d_N and d_S in this region would be almost equal. If overdominant selection operates in the S-RNase gene, d_N would be higher than d_S in the region responsible for the discrimination of self and non-self pollen.

Window analyses of the d_N and d_S of 11 rosaceous S-RNase detected four regions (PS regions) in which d_N was in excess of d_S (Chapter V), evidence that overdominant selection operates in the S-RNase gene if these PS regions are responsible for discrimination between self and non-self pollen. Because four amino acid substitutions in the PS1 and PS2 regions of the *Solanum chacoense* S_{11}- and S_{13}-RNases are necessary for discrimination between S_{11}- and S_{13}-pollen (Matton *et al.*, 1997), overdominant selection must operate, at least in the

PS1 or PS2 region. In the PS1 region, d_S is higher than in other regions and is roughly equal to d_N in the PS1 region. Interestingly, the number of base substitutions in the intron inserted in the PS1 region is as high as the d_N and d_S in the PS1 region. The number of base substitutions in the PS1 region that has the intron was higher than in the other regions of the *P. pyrifolia* genome, indicative that a high mutation rate, gene conversion, or interlocus genetic exchange in this region also may contribute to the high degree of S-RNase polymorphism. But, because interlocus exchange in the S-RNase gene causes change in the pollen *S*-phenotype, simple gene exchange is unlikely. In fact, the phylogenetic trees of exon 1 (from the *N*-terminus to the intron), exon 2 (from the intron to the *C*-terminus), and the whole of S-RNase have the same branching pattern (data not shown). Double recombination (gene conversion) may occur in this region. Nucleotide sequence analysis of many S-RNase genes is necessary to determine the driving force of the polymorphism of the rosaceous S-RNase.

Application to horticulture

Japanese pear is an important commercial fruit tree crop, and breeders have been challenged to produce a new disease-resistant cultivar that bears excellent fruit (Kajiura and Sato, 1990). To avoid using artificial pollination, which requires many laborers, they have tried to breed a self-compatible cultivar(s) of excellent quality. The *S*-genotype of the pear cultivar is an important factor for crossing and breeding, and *S*-genotypes of Japanese pear cultivars have been determined by crossing experiments, but it takes five to ten years to determine the *S*-genotype of a cultivar by this method. Crossing experiments sometimes produce ambiguous results, because the SI of the Japanese pear is not a clear phenomenon. In fact, more than 70 % of its fructification is defined as 'compatible' in crossing experiments, less than 30 % as 'incompatible'. In

addition, these experiments are markedly affected by environmental and physiological factors.

To overcome these weaknesses, I developed a molecular biological method by which to identify the *S*-allele of Japanese pear cultivars that is based on the structural information obtained in this study. The 2D-PAGE method established in chapter I was used to determine the *S*-genotype of the cultivar 'Hosui', often the parent used to breed excellent cultivars but whose *S*-genotype could not be established from crossing experiments (Ishimizu *et al.*, 1998a). The method is rapid (three days) and reliable, but it has not spread widely among breeders because its use is complicated and requires a skillful technique and mature flowers.

I also used the PCR and *S*-allele-specific digestion with restriction endonucleases to develop a method for identifying *S*-alleles (Ishimizu *et al.*, in preparation). Using the two nucleotide sequences conserved in S-RNase genes, I amplified specific fragments of the S-RNase gene from DNA prepared from 0.1 g of young pear leaves. The amplified fragments were digested with each of the six respective *S*-allele-specific restriction endonucleases. The unknown *S*-genotypes of nine cultivars of Japanese pear could be determined by this procedure. As an application of the method, self-compatible varieties could be easily and rapidly selected from offspring of the self-compatible cultivar 'Osa-Nijisseiki' during the early stage of screening.

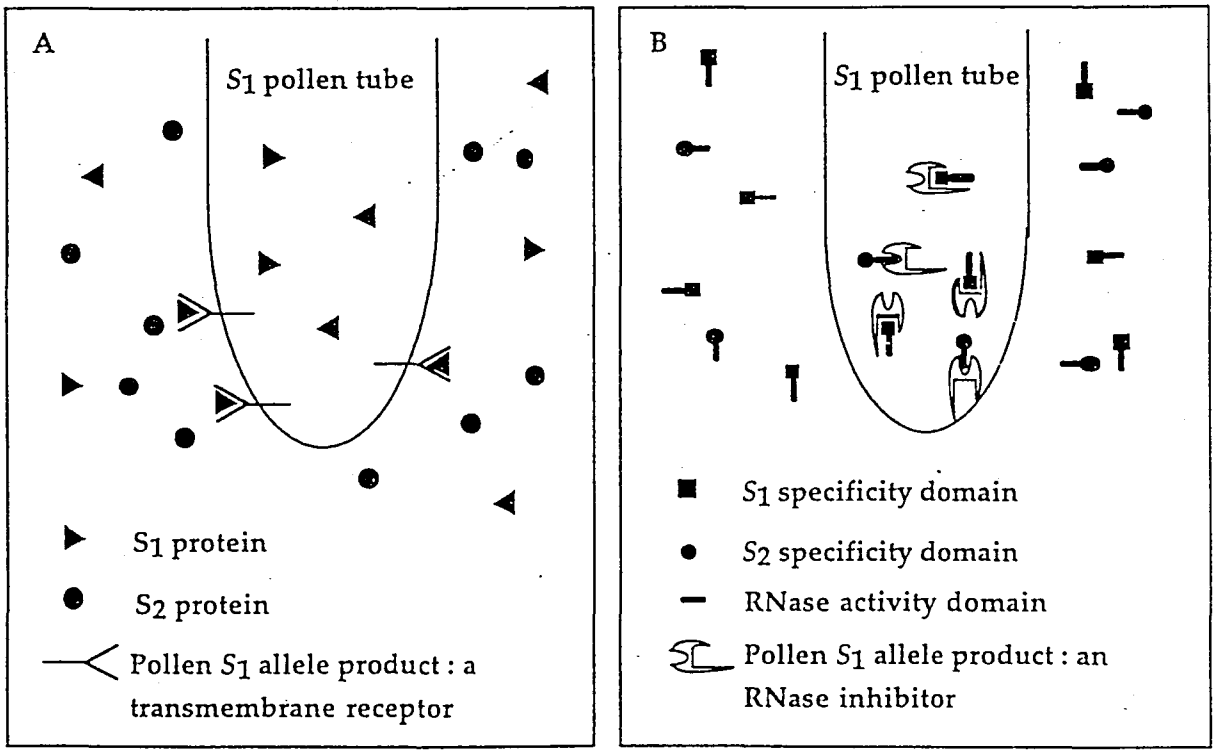


Figure 3. **Two molecular mechanism models for RNase-based GSI (reprinted from Kao and McCubbin, 1996)**. The S_1 pollen tube grows in the transmitting tissue of the S_1S_2 style in both model. (A) *S*-allele-specific uptake model. Specific uptake of S_1-RNase into the S_1 pollen tube results in inhibition of pollen tube growth. (B) RNase inhibitor models. S-RNase enters the S_1 pollen tube non-specifically, but only S_1-RNase functions as an RNase, resulting in inhibition of pollen tube growth.

Summary

I clarified that the GSI-associated S-proteins expressed in pistils of *P. pyrifolia* of the Rosaceae are ribonucleases (S-RNases) also determined and characterized their structures in detail. How S-RNase discriminates self and non-self pollen was discussed on the basis of these S-RNase structural characteristics.

S-allele-specific proteins which had ribonuclease activity obtained from the style extract of *P. pyrifolia,* were identified by 2D-PAGE. S_2- and S_4-RNases were expressed in the pistils of the cultivar 'Nijisseiki' (S_2S_4) when it acquired and intensified SI. 'Osa-Nijisseiki', a self-compatible mutant of 'Nijisseiki', produced S_2-RNase, but not S_4-RNase. These findings suggest that *P. pyrifolia* of the Rosaceae has a GSI system in which an S-RNase acts (Chapter I).

The primary structures of the seven S-RNases identified were determined. Alignment of their amino acid sequences with those of the other rosaceous S-RNases showed that 76 conserved amino acid residues were spread throughout the sequence but were absent from the region from the 51st to 66th residue, designated the HV region. In the *P. pyrifolia* S_3- and S_5-RNase pair (95.5 % homology), amino acid substitutions were present only in the 21st to 90th residue region which includes the HV region. This suggests that this restricted region is responsible for discriminating between S_3- and S_5-pollen (Chapter II).

The locations of the disulfide bonds and free cysteine residues in the S-RNases were determined. Proteins *S*-pyridylethylated at an acidic pH were digested with API at pH 6.5 then analyzed by LC/ESI-MS. The *N. alata* S_6-RNase has two free cysteine residues, Cys77 and Cys95, and four disulfide bonds at

Cys16-Cys21, Cys45-Cys94, Cys153-Cys182 and Cys165-Cys176. The *P. pyrifolia* S_4-RNase also has four disulfide bonds that correspond to those of the *N. alata* S_6-RNase. Eight cysteines that form disulfide bonds are conserved in almost all the S-RNases, which suggests that these cross-links are important in the folding or stabilizing of the tertiary structures of the S-RNases (Chapter III).

The *N*-glycans structures of the *P. pyrifolia* S-RNases were determined by the combination of chromatographic analysis of the PA-sugar chains and LC/ESI-MS analysis. The S-RNases have various types of sugar chains, including such short type one as *N*-acetylglucosamine and chitobiose. Almost all the sites in the seven S-RNases are glycosylated heterogeneously. The S_3- and S_5-RNases, a highly homologous pair, have the same *N*-glycosylation sites and very similar *N*-glycans at each site, indicative that the amino acid moiety, not the carbohydrate, is associated with the discrimination of the *S*-alleles of pollen (Chapter IV).

Window analysis of the d_S and d_N in the rosaceous S-RNases detected four regions with an excess of d_N over d_S, in which positive selection may operate (PS regions). The topology of the predicted secondary structure of these S-RNases is very similar to that of the known tertiary structure of fungal RNase Rh. When the S-RNases, sequences were aligned with the sequence of RNase Rh, the four PS regions corresponded to two surface sites on the tertiary structure of RNase Rh. Because of the similarity of the S-RNases to other proteins responsible for recognition, the two sites formed by the PS regions probably function in *S*-allele-specific discrimination (Chapter V).

Overdominant selection is considered to operate in the PS1 region because this region of the *Solanum chacoence* S_{11}- and S_{13}-RNases is responsible for the discrimination of S_{11} and S_{13} pollen (Matton *et al.*, *Plant Cell* **9**, 1757-1766, 1997). A high mutation rate or interlocus genetic exchange also may operate in the PS1 region because d_S was as high as d_N in this region.

References

Ai, Y., Kron, E., Kao, T.-h. (1991) *S*-alleles are retained and expressed in a self-compatible cultivar of *Petunia hybrida*. *Mol. Gen. Genet.* **230**, 353-358.

Ai, Y., Singh, A., Coleman, C.E., Ioerger, T.R., Kheyr-Pour, A., Kao, T-h. (1990) Self-incompatibility in *Petunia inflata*: isolation and characterization of cDNAs encoding three *S*-allele-associated proteins. *Sex. Plant Reprod.* **3**, 130-138.

Anderson, M.A., Cornish, E.C., Mau, S.-L., Williams, E.G., Hoggart, R., Atkinson, A., Bonig, I., Grego, B., Simpson, R., Roche, P.J., Haley, J.D., Penschow, J.D., Niall, H.D., Tregear, G.W., Coghlan, J.P., Crawford, R.J., Clarke, A.E. (1986) Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. *Nature* **321**, 38-44.

Anderson, M.A., McFadden, G.I., Bernatzky, R., Atkinson, A., Orpin, T., Dedman, H., Tregear, G., Fernley, R., Clarke, A.E. (1989) Sequence variability of three alleles of the self-incompatibility gene of *Nicotiana alata*. *Plant Cell* **1**, 483-491.

Banno, K., Kumashiro, K., Tateishi, S., Takamizawa, M., Kimura, Y., Tokuhisa, T., Tamura, F., Tanabe, K. (1993) Breeding of intergeneric hybrids between Japanese pear and the apple. *J. Japan. Soc. Hort. Sci.* **62** (Suppl. 1), 138-139 (in Japanese).

Bariola, P.A., Howard, C.J., Taylar, C.B., Verburg, M.T., Jaglan, V.D., Green, P.J. (1994) The *Arabidopsis* ribonuclease gene *RNS1* is tightly controlled in response to phosphate limitation. *Plant J.* **6**, 673-685.

Broothaerts, W., Jansserns, G.A., Proost, P., Broekaert, W.F. (1995) cDNA cloning and molecular analysis of two self-incompatibility alleles from apple. *Plant Mol. Biol.* **27**, 499-511.

Castelli, J.C., Hassel, B.A., Wood, K.A., Li, X.L., Amemiya, K., Dalakas, M.C., Torrence, P.F., Youle, R.J. (1997) A study of the interferon antiviral mechanism: apoptosis activation by the 2-5A system. *J. Exp. Med.* **186**, 967-972.

Challice, J.S. (1974) Rosaceae chemotaxonomy and the origins of the Pomoidae. *Bot. J. Linn. Soc.* **69**, 239-259.

Clark, A.G., Kao, T-h. (1991) Excess nonsynonymous substitution at shared polymorphic sites among self-incompatibility alleles of Solanaceae. *Proc. Natl. Acad. Sci. USA* **88**, 9823-9827.

Clark, K.R., Sims, T.L. (1994) The S-ribonuclease gene of *Petunia hybrida* is expressed in nonstylar tissue, including immature anthers. *Plant Physiol.* **106**, 25-36.

Cornish, E.C., Pettitt, J.M., Bonig, I., Clarke, A.E. (1987) Developmentally controlled expression of a gene associated with self-incompatibility in *Nicotiana alata*. *Nature* **326**, 99-102.

Crane, M.B. (1925) Self-sterility and cross incompatibility in plums and cherries. *J. Genet.* **15**, 301-322.

Crestfield, A.M., Moore, S., Stein, W.H. (1963) The preparation and enzymatic hydrolysis of reduced and *S*-carboxymethylated proteins. *J. Biol. Chem.* **238**, 622-627.

Darwin, C. (1876) Effect of cross and self fertilisation in the vegetable kingdom. John Murray, London.

Dodds, P.N., Bönig, I., Du, H., Rödin, J., Anderson, M.A., Clarke, A.E. (1993) S-RNase gene of *Nicotiana alata* is expressed in developing pollen. *Plant Cell* **5**, 1771-1782.

Dodds, P.N., Clarke, A.E., Newbigin, E. (1996) A molecular perspective on pollination in flowering plants. *Cell* **85**, 141-144.

Endo, T., Ikeo, K., Gojobori, T. (1996) Large-scale search for genes on which positive selection may operate. *Mol. Biol. Evol.* **13**, 685-690.

Endo, Y., Huber, P.W., Wool, I.G. (1983) The ribonuclease activity of the cytotoxin α-sarcine: the characterstics of the enzymatic activity of α-sarcine with ribosomes and ribonucleis acids as substrates. *J. Biol. Chem.* **258**, 2662-2667.

Faye, L., Strum, A., Bollini, R., Chrispeels, M.J. (1986) The position of the oligosaccharide side-chains of phytohemagglutinin and their accessibility to glycosidases determines their subsequent processing in the Golgi. *Eur. J. Biochem.* **158**, 655-661.

Fernandez, J., Andrews, L., Mische, S.M. (1994) An improved procedure for enzymatic digestion of polyvinylidene difluoride-bound proteins for internal sequence analysis. *Anal. Biochem.* **218**, 112-117.

Foote, H.C.C., Ride, J.P., Franklin-Tong, V.E., Walker, E.A., Lawrence, M.J., Franklin, F.C.H. (1994) Cloning and expression of a distinctive class of self-incompatibility (*S*) gene from *Papaver rhoeas* L.. *Proc. Natl. Acad. Sci. USA* **91**, 2265-2269.

Franklin-Tong, V.E., Atwal, K.K., Howell, E.C., Lawrence, M.J., Franklin, F.C.H. (1991) Self-incompatibility in *Papaver rhoeas*: there is no evidence for the

involvement of stigmatic ribonuclease activity. *Plant Cell Environ.* **14**, 423-429.

Friedman, M., Krull, L.H., Cavins J.F. (1970) The chromatographic determination of cystine and cysteine residues in proteins as *S*-ß-(4-pyridylethyl)cysteine. *J. Biol. Chem.* **245**, 3868-3871.

Furuta, O., Imai, T., Miyoshi, T., Yatsumi, N., Ueki, S., Hayashi, S., Hiragi, S. (1980) Properties of Japanese pear 'Osa-Nijisseiki'. *J. Japan. Soc. Hort. Sci.* 49 (Suppl. 2) pp.70-71 (in Japanese).

Green, P.J. (1994) The ribonucleases of higher plants. *Annu. Rev. Plant Physiol. Mol. Biol.* **45**, 421-445.

Hase S. (1994) High-performance liquid chromatography of pyridylaminated saccharides. *Methods in Enzymol.* **230**, 225-237.

Hase, S., Koyama, S., Daiyasu, H., Takemoto, H., Hara, S., Kobayashi, Y., Kyogoku, Y., Ikenaka, T. (1986) Structure of a sugar chain of a protease inhibitor isolated from Barbados pride (*Caesalpinia pulcherrima* SW.) seeds. *J. Biochem.* **100**, 1-12.

Hase S., Ikenaka, K., Mikoshiba, K., Ikenaka, T. (1988) Analysis of tissue glycoprotein sugar chains by two-dimensional high-performance liquid chromatographic mapping. *J. Chromatogr.* **434**, 51-60.

Hase, S., Hatanaka, K., Ochiai, K., Shimizu, H. (1992) Improved method for the compornent sugar analysis of glycoproteins by pyridylamino sugars purified with immobilized boronic acid. *Biosci. Biotech. Biochem.* **56**, 1676-1677.

von Heijne, G. (1986) A new method for predicting signal sequence cleavage site. *Nucl. Acids Res.* **14**, 4683-4690.

Hinata, K., Watanabe, M., Yamakawa, S., Satta, Y., Isogai, A. (1995) Evolutionary aspects of the *S*-related genes of the Brassica self-incompatibility system: synonymous and nonsynonymous base substitutions. *Genetics* **140**, 1099-1104.

Hirano, H., Watanabe, T. (1990) Microsequencing of proteins electrotransferred onto immobilizing matrices from polyacrylamide gel electrophoresis. *Electrophoresis* **11**, 573-580.

Hiratsuka, S. (1992) Detection and inheritance of a stylar protein associated with a self-incompatibility genotype of Japanese pear. *Euphytica* **61**, 55-59.

Hiratsuka, S., Hirata, N., Tezuka, T., Yamamoto, Y. (1985) Self-incompatibility reaction of Japanese pear in various stages of floral development. *J. Japan. Soc. Hort. Sci.* **54**, 9-14.

Horiuchi, H., Yanai, K., Takagi, M., Yano, K., Wakabayashi, E., Sanda, A., Mine, S., Ohgi, K., Irie, M. (1988) Primary structure of a base non-specific ribonuclease from *Rhizopus niveus*. *J. Biochem.* **103**, 408-418.

Huang, S., Lee. H.S., Karunanaadaa, B., Kao, T-h. (1994) Ribonuclease activity of *Petunia inflata* S proteins is essential for rejection of self-pollen. *Plant Cell* **6**, 1021-1028.

Hughes, A.L., Nei, M. (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* **335**, 167-170.

Hughes, A.L., Nei, M. (1989) Nucleotide substitution at major histocompatibility complex class II loci: Evidence for overdominant slection. *Proc. Natl. Acad. Sci. USA* **86**, 958-962.

Ioerger, T.R., Clarke, A.G., Kao, T-h. (1990) Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proc. Natl. Acd. Sci. USA* **87**, 9732-9735.

Ioerger, T.R., Gohlke, J.R., Xu, B., Kao, T-h. (1991) Primary structural features of the self-incompatibility protein in Solanaceae. *Sex. Plant Reprod.* **4**, 81-87.

Ishimizu, T., Miyagi, M., Norioka, S., Liu, Y-H., Clarke, A.E., Sakiyama, F. (1995) Identification of histidine 31 and cysteine 95 in the active site of self-incompatibility associated S_6-RNase in *Nicotiana alata*. *J. Biochem.* **118**, 1007-1013.

Ishimizu, T., Sato, Y., Saito, T., Yoshimura, Y., Norioka, S., Nakanishi, T., Sakiyama, F. (1996a) Identification and partial amino acid sequences of seven *S*-RNases associated with self-incompatibility of the Japanese pear, *Pyrus pyriforia* Nakai. *J. Biochem.* **120**, 326-334.

Ishimizu, T., Norioka, S., Kanai, M., Clarke, A.E., Sakiyama, F. (1996b) Location of cysteine and cystine residues in S-ribonucleases associated with gametophytic self-incompatibility. *Eur. J. Biochem.* **242**, 627-635.

Ishimizu, T., Norioka, S., Nakanishi, T., Sakiyama, F. (1998a) *S*-genotype of Japanese pear 'Hosui'. *J. Japan. Soc. Hort. Sci.* **67**, 35-38.

Ishimizu, T., Shinkawa, T., Sakiyama, F., Norioka, S. (1998b) Primary structural features of rosaceous S-RNases associated with gametophytic self-incompatibility. *Plant Mol. Biol.* in press.

Jahnen, W., Batterham, M.P., Clarke, A.E., Moritz, R.L., Simpson, R.J. (1989) Identification, isolation, and *N*-terminal sequencing of style glycoproteins associated with self-incompatibility in *Nicotiana alata*. *Plant Cell* **1**, 493-499.

Janssens, G.A., Goderis, I.J., Broekaert, W.F., Broothaerts, W. (1995) A molecular method for *S*-allele identification in apple based on allele-specific PCR. *Theor. Appl. Genet.* **91**, 691-698.

Jost, W., Bak, H., Glund, K., Terpstra, P., Beintema, J.J. (1991) Amino acid sequence of an extracellular, phosphate-stavation-induced ribonuclease from cultured tomato (*Lycopersicon esculentum*)cells. *Eur. J. Biochem.* **198**, 1-6.

Kämper, J., Reichmann, M., Romeis, T., Bölker, M., Kahmann, R. (1995) Multiallelic recognition: nonself-dependent dimerization of the bE and bW homeodomein proteins in *Ustilago maydis*. Cell **81**, 73-83.

Kajiura, I., Sato, Y. (1990) Recent progress in Japanese pear (*Pyrus pyrifolia* Nakai) breeding and descriptions of cultivars based on literature review. Bull. Fruit Tree Res. Stn. Extra No.1. (In Japanese with English summary).

Kao, T-h., McCubbin, A.G. (1996) How flowering plants discriminate between self and non-self pollen to prevent inbreeding. *Proc. Natl. Acad. Sci. USA* **93**, 12059-12065.

Kao, T-h., McCubbin, A.G., Verica, J.A., Wang, X., Skirpan, A.L., Dowd, P.E. (1997) The 5th international congress of plant molecular biology.

Karunanandaa, B., Huang, S., Kao, T-h. (1994) Carbohydrate moiety of the *Petunia inflata* S_3-protein is not required for self-incompatibility interaction between pollen and pistil. *Plant Cell* **6**, 1933-1940.

Kawata, Y., Sakiyama, F., Tamaoki, H. (1988) Amino-acid sequence of ribonuclease T_2 from *Aspergillus oryzae*. *Eur. J. Biochem.* **176**, 683-697.

Kawata, Y., Sakiyama, F., Hayashi, F., Kyogoku, Y. (1989) Identification of two essential histidine residues of ribonuclease T_2 from *Aspergillus oryzae*. *Eur. J. Biochem.* **187**, 255-262.

Kikuchi, A. (1929) Investigations in 1927 and 1928. 1. Paterclinical incompatibility in the Japanese pear. *J. Okitsu Hort. Soc.* **24**, 1-6 (in Japanese).

Kimura, Y., Hase, S., Kobayashi, Y., Kyogoku, Y., Funatsu, G., Ikenaka, T. (1987) Possible pathway for the processing of sugar chains containing xylose in plant glycoproteins deduced on structural analyses of sugar from *Ricinus communis* Lectins. *J. Biochem.* **101**, 1051-1054.

Kirch, H.H., Uhrig, H., Lottspeich, F., Salamini, F., Thompson, R.D. (1989) Characterization of proteins associated with self-incompatibility in *Solanum tuberosum*. *Theor. Appl. Genet.* **78**, 581-588.

Kiyozumi, D. Identification of pollen specific and polymorphic proteins genetically linked to self-incompatibility locus in Japanese pear. Master's thesis (Osaka university).

Kuraya, N., Hase, S. (1992) Release of *O*-linked sugar chains from glycoproteins with anhydrous hydrazine and pyridylamination of the sugar chains with improved reaction conditions. *J. Biochem.* **112**, 122-126.

Kurihara, H., Nonaka, T., Mitsui, Y., Ohgi, K., Irie, M., Nakamura, K.T. (1996) The crystal structure of ribonuclease Rh from *Rhizopus niveus* at 2.0 Å resolution. *J. Mol. Biol.* **255**, 310-320.

Kusaba, M., Nishio, T., Satta, Y., Hinata, K., Ockendon, D. (1997) Striking sequence smilarity in inter- and intra-specific comparisons of class I *SLG* alleles from *Brassica oleracea* and *Brassica campestris*: implications for the evolution and recognition mechanism. *Proc. Natl. Acad. Sci. USA* **94**, 7673-7678.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Latimar, L.P. (1937) Self- and cross-pollination in the McIntosh apple and some of its hybrids. *Proc. Am. Soc. Hortic. Sci.* **34**, 19-21.

Li, X., Nield, J., Hayman, D., Langridge, P. (1994) Cloning a putative self-incompatibility gene from the pollen of the grass *Phalaris coerulescens*. *Plant Cell* **6**, 1923-1932.

Li, X., Nield, J., Hyman, D., Langridge, P. (1996) A self-fertile mutant of *Phalaris* produces an S protein with reduced thioredoxin activity. *Plant J.* **10**, 505-513.

Li, B., Kusaba, M., Li, J.-h., Newbigin, Ed. (1997) The 5th international congress of plant molecular biology.

Lee, H.S., Huang, S., Kao, T-h. (1994) S proteins control rejection of incompatible pollen in *Petunia inflata*. *Nature* **367**, 560-563.

Löffeler, A., Glund, K., Irie, M. (1993) Amino acid sequence of an intracellular, phosphate-starvation-induced ribonuclease from cultured tomato (*Lycopersicon esculentum*) cells. *Eur. J. Biochem.* **214**, 627-633.

Lush, M., Clarke, A.E. (1997) Observation of pollen tube growth in *Nicotiana alata* and their implications for the mechanism of self-incompatibility. *Sex. Plant Reprod.* **10**, 27-35.

Lyons, E.J. (1997) Sex and synergism. *Nature* **390**, 19-21.

Machida, Y. (1972) Morphological and physiological characteristics of Japanese and Chinese pear. In *Dictionary of pomology* (Sato K., Mori, H., Matsui, O., Kitajima, H., Chiba, T. eds.) pp.527-529. Yokendo, Tokyo.

Machida, Y., Sato, Y., Kozaki, I., Seike, K. (1982) *S*-genotype of several cultivars of Japanese pear and the question of the parents of 'Hosui'. *J. Jpn. Soc. Hort. Sci.* **51** (Suppl. 2), 58-59 (in Japanese).

Maekawa, K., Tsunasawa, S., Dibo, G., Sakiyama, F. (1991) Primary structure of nuclease P1 from *Penicillium citrinum*. *Eur. J. Biochem.* **200**, 651-661.

Makino, Y., Kuraya, N., Omichi, K., Hase, S. (1996) Classification of sugar chains of glycoproteins by analyzing reducing end oligosaccharides obtained by partial acid hydrolysis. *Anal. Biochem.* **238**, 54-59.

Maruyama, T., Nei, M. (1981) Genetic variability maintained by mutation and overdominant selection in finite populations. *Genetics* **8**, 444-459.

Matsuo, Y., Yamada, A., Tsukamoto, K., Tamura, H., Ikezawa, H., Nakamura, H., Nishikawa, K. (1996) A distant evolutionary relationship between bacterial sphingomyelinase and mammalian DNase I. *Prot. Sci.* **5**, 2459-2467.

Matsushita, M., Watanabe, M., Yamakawa, S., Takayama, S., Isogai, A., Hinata, K. (1996) The SLGs corresponding to the same S^{24}-haplotype are perfectly conserved in three different self-incompatible *Brassica campestris* L.. *Genes Genet. Syst.* **71**, 255-258.

Matton, D.P., Maes, O., Laublin, G., Xike, Q., Bertrand, C., Morse, D., Cappadocia, M. (1997) Hypervariable domeins of self-incompatibility RNases mediate allele-specific pollen recgnition. *Plant Cell* **9**, 1757-1766.

Mau, S-L., Raff, J., Clarke. A.E. (1982) Isolation and partial characterization of components of *Prunus avium* L. styles, including an antigenic glycoprotein associated with a self-incompatibility genotype. *Planta* **156**, 505-516.

Mau, S-L., Williams, E.G., Atkinson, M.A., Cornish, E.C., Grego, B., Simpson, R.J., Kheyr-Pour, A., Clarke, A.E. (1986) Style proteins of a wild tomato (*Lycopersicon peruvianum*) associated with expression of self-incompatibility. *Planta* **169**, 184-191.

McClure, B.A., Haring, V., Ebert, P.R., Anderson, M.A., Simpson, R.J., Sakiyama, F., Clarke, A.E. (1989) Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. *Nature* **342**, 955-957.

McClure, B.A., Gray, J.E., Anderson, M.A., Clarke, A.E. (1990) Self-incompatibility in *Nicotiana alata* involves degradation of pollen rRNA. *Nature* **347**, 757-760.

McCubbin, A.G., Chung, Y-Y., Kao, T-h. (1997) A mutant S_3 RNase of *Petunia inflata* lacking RNase activity has an allele-specific dominant negative effect on self-incompatibility interactions. *Plant Cell* **9**, 85-95.

McReynolds, L., O'Malley, B.W., Nisbet, A.D., Fothergill, J.E., Givol, D., Fields, S., Robertson, M., Brownlee, G.G. (1978) Sequence of chicken ovalbumin mRNA. *Nature* **273**, 723-728.

Murfett, J., Atherton, T.L., Beiquan, M., Gasser, C.S., McClure, B.A. (1994) S-RNase expressed in transgenic *Nicotiana* causes *S*-allele-specific pollen rejection. *Nature* **367**, 563-566.

Nakamura, K.T., Ishikawa, N., Hamashima, M., Kurihara, H., Nonaka, T., Mitsui, Y., Ohgi, K., Irie, M. (1993) The 3rd international meeting on ribonuclease, chemistry, biology, biotechnology.

Nakanishi, T., Yamazaki, T., Funadera, K., Tomonaga, H., Ozaki, T., Kawai, Y., Ichii, T., Satoh, Y., Kurihara, A. (1992) Isoelectric focusing analysis of stylar proteins associated with self-incompatibility alleles in Japanese pear. *J. Japan. Soc. Hort. Sci.* **61**, 239-248.

Nasrallah, J.B., Kao, T.-h., Goldberg, M.L., Nasrallah, M.E. (1985) A cDNA clone encoding an *S*-locus-specific glycoprotein from *Brassica oleracea*. *Nature*, **318**, 263-267.

Nasrallah, J.B., Kao, T.-h., Chen, C.-h., Goldberg, M.L., Nasrallah, M.E. (1987) Amino acid sequence of glycoproteins encoded by three alleles of the *S*-locus of *Brassica oleracea*. *Nature*, **326**, 617-619.

Nasrallah, J.B. (1997) Evolution of the *Brassica* self-incompatibility locus: A look into *S*-locus gene polymorphisms. *Proc. Natl. Acad. Sci. USA* **94**, 9516-9519.

de Nettancourt, D. (1977) Incompatibility in Angiosperms. Springer, NewYork.

Nei, M., Gojobori, T. (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**, 428-426.

Norioka, N., Ohnishi, Y., Norioka, S., Ishimizu, T., Nakanishi, T., Sakiyama, F. (1995) Nucleotide sequences of cDNAs encoding S_2- and S_4-RNases (D49527 and D49528 for EMBL) from Japanese pear (*Pyrus pyrifolia* Nakai) (PGR95-020). *Plant Physiol.* **108**, 1343.

Norioka, N., Norioka, S., Ohnishi, Y., Ishimizu, T., Oneyama, C., Nakanishi, T., Sakiyama, F. (1996) Molecular cloning and nucleotide sequences of cDNAs encoding *S*-allele specific stylar RNases in a self-incompatible cultivar and

its self-compatible mutant of Japanese pear, *Pyrus pyrifolia* Nakai. *J. Biochem.* **120**, 335-345.

O'Farrell, P.H. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**, 4007-4021.

O'Farrell, P.Z., Goodman, H.M., O'Farrell, P.H. (1977) High resolution two-dimensional electrophoresis of basic as well as acidic proteins. *Cell* **12**, 1133-1142.

Oku, H., Hase, S., Ikenaka, T. (1990) Separation of oligomannose-type sugar chains having one to five mannose residues by high-performance pyridylamino derivatives. *Anal. Biochem.* **185**, 331-334.

Ota, M., Nishikawa, K. (1997) Assessment of pseudo-energy potentials by the best-five test: a new use of the three-dimensional profiles of proteins. *Prot. Eng.* **10**, 339-351.

Oxley, D., Bacic, A. (1995) Microheterogeneity of *N*-glycosylation on a stylar self-incompatibility glycoprotein of *Nicotiana alata*. *Glycobiology* **5**, 517-523.

Oxley, D., Munro, S.L.A., Craik, D.J., Bacic, A. (1996) Structure of *N*-glycan on the S_3- and S_6-allele stylar self-incompatibility ribonucleases of *Nicotiana alata*. *Glycobiology* **6**, 611-618.

Rich, A., Nordheim, A., Wang, A. H-j. (1984) The chemistry and biology of left-handed Z-DNA. *Ann. Rev. Biochem.* **53**, 791-846.

Rost, B., Sander, C. (1993) Prediction of protein secondary structure at better than 70 % accuracy. *J. Mol. Biol.* **232**, 584-599.

Royo, J., Kowyama, Y., Clarke, A.E. (1994) Cloning and nucleotide sequence of two S-RNases from *Lycopersicon peruvianum* (L.) Mill. Plant Physiol. **105**, 751-752.

Rudd, P.M., Joao, H.C., Coghill, E., Fiten, P., Saunders, M.R., Opdenakker, G., Dwek, R.A. (1994) Glycoforms modify the dynamic stability and functional activity of an enzyme. *Biochemistry* **33**, 17-22.

Saba-El-Leil, M.K., Rivard, S., Morse, D., Cappadocia, M. (1994) The S11 and S13 self incompatibility alleles in *Solanum chacoense* Bit. are remarkably similar. *Plant Mol. Biol.* **24**, 571-583.

Saitou, N., Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425.

Sassa, H., Hirano, H., Ikehashi, H. (1992) Self-incompatibility-related RNases in styles of Japanese pear (*Pyrus serotina* Rehd). *Plant Cell Physiol.* **33**, 811-814.

Sassa, H., Hirano, H., Ikehashi, H. (1993) Identification and characterization of stylar glycoproteins associated with self-incompatibility genes of Japanese pear, *Pyrus serotina* Rehd.. *Mol. Gen. Genet.* **241**, 17-25.

Sassa, H., Mase, N., Hirano, H., Ikehashi, H. (1994) Identification of self-incompatibility-related glycoproteins in styles of apple (*Malus x domestica*). *Theor. Appl. Genet.* **89**, 201-205.

Sassa, H., Nishio, T., Kowyama, Y., Hirano, H., Koba, T., Ikehashi, H. (1996) Self-incompatibility (*S*) alleles of the Rosaceae encode members of a distinct class of the T_2 / *S* ribonuclease superfamily. *Mol. Gen. Genet.* **250**, 547-557.

Sassa, H., Hirano, H. (1997) Nucleotide sequence of a cDNA encoding S_5-RNase (Accession No. D88282) from Japanese pear (*Pyrus serotina*) (PGR97-007). *Plant Physiol.* **113**, 306.

Sassa, H., Hirano, H., Nishio, T., Koba, T. (1997) Style-specific self-incompatible mutation caused by deletion of the S-RNase gene in Japanese pear (*Pyrus serotina*). *Plant J.* **12**, 223-227.

Sato, Y. (1992) Breeding of self-compatible Japanese pear. *Abst. the Symposium of Japan. Soc. Hort. Sci.* pp.12-22 (in Japanese).

Sato, Y. (1993) Breeding of self-compatible Japanese pear. In Techniques on Gene Diagnosis and Breeding in Fruit Trees, Hayashi, T., *et al.*, eds (Tokyo: FTRS / Japan), pp.241-247.

Shapiro, R., Fett, J.W., Strydom, D.J., Vallee, B.L. (1986) Isolation and characterization of a human colon carcinoma-serected enzyme with pancreatic ribonuclease like activity. *Biochemistry* **25**, 7255-7264.

Shimura, I., Ito, Y., Seike, K. (1983) Intergeneric hybrid between Japanese pear and quince. *J. Japan. Soc. Hort. Sci.* **52**, 243-249. (In Japanese with English abstract).

Shinkawa, T. (1997) Primary structural analysis and enzymatic characterization of stylar S-RNase of *Pyrus pyrifolia*. Master's thesis (Osaka University) (in Japanese).

Singh, A., Ai, Y., Kao, T-h. (1991) Characterization of ribonuclease activity of three S-allele-associated proteins of *Petunia inflata*. *Plant Physiol.* **96**, 61-68 .

Stein, J.C., Howlett, B., Boyes, D.C., Nasrallah, M.E., Nasarallah, J.B. (1991) Molecular cloning of a putative receptor kinase gene encoded at the self-incompatibility locus of *Brassica oleracea*. *Proc. Natl. Acad. Sci. USA* **8**, 8816-8820.

Suzuki, J., Kondo, A., Kato, I., Hase, S., Ikenaka, T. (1991) Analysis by high-performance anion-exchange chromatography of compornent sugars as their fluorescent pyridylamino derivatives. *Agric. Biol. Chem.* **55**, 283-284.

Takayama, S., Isogai, A., Tsukamoto, C., Ueda, Y., Hinata, K., Okazaki, K., Suzuki, A. (1987) Sequences of S-glycoproteins, products of the *Brassica campestris* self-incompatibility locus. *Nature*, **326**, 102-104.

Takuma, S. (1996) Identification and purification of stylar RNases associated with self-incompatibility from the *P. pyrifolia* cultivar 'Nijisseiki' (S_2S_4). Master's thesis (Osaka University) (in Japanese).

Tao, R., Yamane, H., Sassa, H., Mori, H., Gradziel. T.M., Dandekar, A.M., Sugiura, A. (1997) Identification of stylar RNases associated with gametophytic self-incompatibility in almond (*Prunus dulicis*). *Plant Cell Physiol.* **38**, 304-311.

Taylor, C.B., Bariola, P.A., del Cardayré, S.B., Rains, R.T., Green, P.J. (1993) RNS2: A senescence-associated RNase of *Arabidopsis* that diverged from the *S*-RNase before speciation. *Proc. Natl. Acad. Sci. USA.* **90**, 5118-5122.

Terai, O., Sato, Y., Abe, K., Saito, T., Kotobuki, K. (1995) *S*-gene homozygote : useful tool for *S*-genotype analysis in Japanese pear. *Abst. Japan. Soc. Hort. Sci. Autumn Meet.* pp.150-151 (in Japanese).

Terami, H., Torikata, H., Shimazu, Y. (1946) Analysis of the sterility-factors existing in varieties of the Japanese pear (*Pyrus serotina* Rehd. *var. culta* Rehd.). *Studies Hort. Inst. Kyoto Imp. Univ.* **3**, 267-271 (in Japanese with English abstract).

Thompson, E.O.P., Fisher, W.K. (1978) Amino acid sequences containing half-cystine residues in ovalbumin. *Aust. J. Biol. Sci.* **31**, 433-442.

Thompson, R.D., Uhrig, H., Hermsen, J.G., Salamini, F., Kaufmann, H. (1991) Investigation of a self-compatible mutation in *Solanum tuberosum* clones inhibiting *S*-allele activity in pollen differentially. *Mol. Gen. Genet.* **226**, 283-288.

Tomimoto, Y., Nakazaki, T., Ikehashi, H., Ueno, H., Hayashi, R. (1997) Analysis of self incompatibility-related riblnucleases (S-RNases) in two species of pears, *Pyrus communis* and *P. ussuriensis*. *Sci. Hort.* **66**, 159-167.

Tsai, D.S., Lee, H.S., Post, L.C., Kreiling, K.M., Kao, T-h. (1992) Sequence of an S-protein of *Lycopersicon peruvianum* and comparison with other solanaceous S-proteins. *Sex. Plant Reprod.* **5**, 256-263.

Tsunasawa, S., Kondo, J., Sakiyama, F. (1985) Isocratic separation of PTH-amino acids at picomole level by reverse-phase HPLC in the presence of sodium dodecylsulfate. *J. Biochem.* **97**, 701-704.

Walker, E.A., Ride, J.P., Kurup, S., Franklin-Tong, V.E., Lawrence, M.J., Franklin, F.C.H. (1996) Molecular analysis of two functional homologues of the S_3 allele of the *Papaver rhoeas* self-incompatibility gene isolated from different populations. *Plant Mol. Biol.* **30**, 983-994.

Watson, M.E.E. (1984) Copilation of published signal sequences. *Nucl. Acids Res.* **12**, 5145-5164.

Westwood, M.N., Lombard., P.B., Bjornstad, H.O. (1989) Pear on 'Winter Banana' interstem with M. 26 apple rootstock as a compatible combination. *HortScience* **24**, 765-767.

Woods, R.J., Edge, C.J., Dewk, R.A. (1994) Protein surface oligosaccharides and protein function. *Nat. Struct. Biol.* **1**, 499-501.

Wright, S. (1939) The distribution of self-fertility alleles in populations. *Genetics* **24**, 538-552.

Xue, Y., Carpenter, R., Dickinson, H.G., Coen, E.S. (1996) Origin of allelic diversity in antirrhinum *S* locus RNases. *Plant Cell* **8**, 805-814.

Yanagida, K., Ogawa, H., Omichi, K., Hase, S. (1998) Introduction of a new scale into reversed-phase HPLC of pyridylamino sugar chains for structural assignment. *J. Chromatogr.* in press.

Zurek, D.M., Mou, B., Beecher, B., McClure, B. (1997) Exchanging sequence domeins between S-RNases from Nicotiana alata disrupts pollen recognition. *Plant J.* **11**, 797-808.

List of publications

1. Norioka, N., Ohnishi, Y., Norioka, S., Ishimizu, T., Nakanishi, T., and Sakiyama, F. (1995) Nucleotide sequences of cDNAs encoding S_2- and S_4-RNases (D49257 and D49528 for EMBL) from Japanese pear (*Pyrus Pyriforia* Nakai). *Plant Physiol.* **108**, 1343.
2. Ishimizu, T., Sato, Y., Saito, T., Yoshimura, Y., Norioka, S., Nakanishi, T., and Sakiyama, F. (1996) Identification and partial amino acid sequences of seven S-RNases associated with self-incompatibility of Japanese pear, *Pyrus pyrifolia* Nakai. *J. Biochem.* **120**, 326-334.
3. Norioka, N., Norioka, S., Ohnishi, Y., Ishimizu, T., Oneyama, C., Nakanishi, T., and Sakiyama, F. (1996) Molecular cloning and nucleotide sequences of cDNAs encoding S-alleles specific pistil RNases in a self-incompatible cultivar and its self-compatible mutant of Japanese pear (*Pyrus pyrifolia* Nakai). *J. Biochem.* **120**, 335-345.
4. Ishimizu, T., Norioka, S., Kanai, M., Clarke, A.E., and Sakiyama, F. (1996) Location of cysteine and cystine residues in tobacco S_6-RNase and Japanese pear S_4-RNase associated with gametophytic self-incompatibility. *Eur. J. Biochem.* **242**, 627-635.
5. Ishimizu, T., Shinkawa, T., Sakiyama, F., and Norioka, S. (1998) Primary structural features of rosaceous S-RNases associated with gametophytic self-incompatibility. *Plant Mol. Biol.* in press.
6. Ishimizu, T., Endo, T., Gojobori, T., Nakamura, K.T., Sakiyama, F., and Norioka, S. Identification of regions in which positive selection may operate in S-RNase of Maloideae in Rosaceae: Implication for *S*-allele-specific recognition sites in S-RNase. in preparation
7. Ishimizu, T., Mitsukami, Y., Miyagi, M., Shinkawa, T., Natsuka, S., Sakiyama, F., Norioka, S., and Hase, S. Comparison of sites and structures of *N*-glycan

containing a novel structure, a chitobiose, on seven S-RNases associated with self-incompatibility in *Pyrus pyrifolia.* in preparation.

Related papers

1. Ishimizu, T., Miyagi, M., Norioka, S., Liu, Y.-H., Clarke, A.E., and Sakiyama, F. Identification of histidine 31 and cysteine 95 in the active site of self-incompatibility associated S_6-RNase in *Nicotiana alata*. *J. Biochem.* **118**, 1007-1013 (1995)
2. Ishimizu, T., Norioka, S., Nakanishi, T., and Sakiyama, F. *S*-genotype of Japanese pear 'Hosui'. *J. Japan. Soc. Hort. Sci.* **67**, 35-38 (1998).
3. Norioka, S., Oneyama, C., Takuma, S., Ishimizu, T., Nakanishi, T., and Sakiyama, F. Isolation and primary structure of a pear non-S-RNase highly homologous to tomato and tobacco stress-induced RNases. in preparation.
4. Ishimizu, T., Inoue, K., Shimonaka, M., Saito, T., Terai, O., and Norioka, S. PCR-based method for determination of *S*-genotype of Japanese pear. in preparation.
5. Ishimizu, T., Norioka, N., Matsuki, T., Sakiyama, F., and Norioka, S. Nucleotide sequences of intron in seven *Pyrus pyrifolia* S-RNases. in preparation.

Acknowledgements

This dissertation is dedicated to all the people who supported me.

The studies described in this thesis have been carried out under the direction of (Emeritus) Professor Fumio Sakiyama, Institute for Protein Research, Osaka University. I would like to express my sincere gratitude to Professor Fumio Sakiyama for his firm guidance, essential criticism, and kind consideration.

I am deeply indebted to Associate Professor Shigemi Norioka (Institute for Protein Research) for his helpful criticism and comments, various clerical procedures, and for his kindness. Thanks are also due to Professor Saburo Aimoto (Institute for Protein Research) for his critical comments for submitting this dissertation.

I wishes to express my gratitude to all the members of the 'Nashi Group' (our research group for the molecular mechanism of GSI). I wishes to express my gratitude to Professor Tetsu Nakanishi (Faculty of Agriculture, Kobe University) for teaching the physiology of SI of *P. pyrifolia*, Mrs. Yoshiro Mitsukami and Toyohide Shinkawa (Institute for Protein Research) for cooperation in structural analysis of primary structures and sugar chains of S-RNases, Mr. Daiji Kiyozumi (Institute for Protein Research) for exciting discussion, and other members (Drs. Naoko Norioka and Yuko Natsuka, Mses. Chitose Oneyama, Yoshimi Ohnishi, and Satoko Okuhata, Mrs. Seiji Takuma, Toru Matsuki, and Takanori Matsuura of Institute for Protein Research) for their kind cooperation and their friendship.

I also wishes to express my gratitude to the staffs of Division of Protein Chemistry, Institute for Protein Research, Dr. Shaoliang Li for helpful comments, Mses. Yoshiko Yagi and Yumi Yoshimura for amino acid sequence analysis, and Ms. Chizu Sasai for clerical work.

I would like to express my thanks to the members of the Tottori Horticultural Experiment Station including Mrs. Kosuke Inoue, Masahito Shimonaka, Hiroki

Tabira and the members of the Fruit Tree Research Station of the Ministry of Agriculture, Forestry and Fisheries of Japan including Mrs. Yoshihiko Sato, Toshihiro Saito, and Osamu Terai for supplying the flowers of the *P. pyrifolia* cultivars and for teaching the physiology of *P. pyrifolia*.

Furthermore, I would like to express my thanks to Professor Sumihiro Hase (Faculty of Science, Osaka University) for guidance of structural analysis of sugar chains and of the attitude toward scientific research. I also would like to express my thanks to the members of Professor Hase's laboratory including Dr. Shunji Natsuka, Mrs. Yasushi Makino and Kanta Yanagida for their technical assistance, kind cooperation, and friendship.

I am indebted to two skillful researchers, Dr. Masaru Miyagi (Takara Shuzo Co.) and Dr. Michiko Kanai (Thermoquest Co.) for mass spectrometry and substantial discussion.

I am also indebted to Professor Takashi Gojobori, Drs. Toshinori Endo and Yumi Yamaguchi (National Institute of Genetics) for estimating the number of nucleotide substitutions in the S-RNases gene and for their helpful discussions on high degree of polymorphism of S-RNase.

I would like to express thanks to Professor Adrienne E. Clarke (Melbourne University, Australia) for her kindly supply of *N. alata* S_6-RNase and collaboration.

Finally, I am grateful to my parents for their unfailing financial support and great affection.

石水 毅

Takeshi Ishimizu

February, 1998