The EMBO Journal vol.4 no.11 pp 2745-2749, 1985

Amino acid sequence of the 'pathogenesis-related' leaf protein p14 from viroid-infected tomato reveals a new type of structurally unfamiliar proteins

J.Lucas[1,2], A.Camacho Henriquez[1], F.Lottspeich[2], A.Henschen[2] and H.L.Sänger[1]

[1]Abteilung für Viroidforschung, and [2]Abteilung für Proteinchemie, Max-Planck-Institut für Biochemie, D-8033 Planegg-Martinsried bei München, FRG

Communicated by H.L.Sänger

We have established the complete sequence of the 130 amino acid residues of the pathogenesis-related (PR) protein p14 accumulating in tomato leaves infected with the viroid of the spindle tuber disease of potato (PSTV) and partial sequences of the PR protein 1a which accumulates in tobacco mosaic virus-infected tobacco leaves. Both PR proteins are closely related to each other. However, no homology could be found between the sequence of p14 and any of the 3061 published protein sequences compiled in the protein sequence database at present. Thus, p14 represents not only the first completely sequenced PR plant protein but also a new type of structurally unfamiliar proteins whose biological function in the diseased plant remains to be elucidated.

Key words: disease-stimulated protein/pathogenesis-related protein/plant protein/tomato leaf protein

Introduction

Upon infection of tomato plants with either viroids, or two conventional viruses or the leaf mold *Cladosporium fulvum*, a basic leaf protein with an apparent mol. wt. of 14 000 accumulates to become the most abundant acid-extractable tomato leaf protein. This 'pathogenesis-related' (PR) protein, referred to as p14, is also found in trace amounts in healthy, especially aging leaves. Its accumulation in infected, symptom-bearing leaves has been shown to be a general pathophysiological but pathogen-unspecific response of the tomato plant to infection (Camacho Henriquez and Sänger, 1982a, 1982b, 1984). The accumulation of such host-encoded soluble leaf proteins has been first described for tobacco plants infected with tobacco mosaic virus (TMV) (van Loon and van Kammen, 1970; Gianinazzi *et al.*, 1970, 1977). Since their appearance could at first only be related to a pathological condition they were named 'pathogenesis-related' or 'PR' proteins (Antoniw *et al.*, 1980; van Loon *et al.*, 1983), a term which has largely replaced the name 'b' proteins also used in the past (Gianinazzi *et al.*, 1970). Thus far, PR proteins have been detected in leaves of 16 different plant species from which they are preferentially extracted at an acid pH value and the mol. wt. of all of them range between 10 000 and 20 000 (reviewed by van Loon, 1985). It should be noted, however, that the accumulation of proteins with an electrophoretic behaviour similar to the one of the PR proteins can also be induced artificially in healthy plants by plasmolysis (Wagih and Coutts, 1981), by the application of a variety of chemicals (reviewed by van Loon, 1983a, 1983b, 1985), by high concentrations of various plant hormones (van Loon, 1977, 1983a, 1983b; Antoniw *et al.*, 1981) and by culture filtrates from a pathogenic fungus (Maiss and Poehling, 1983) and from a bacterium (Leach *et al.*, 1983). Such proteins even appear spontaneously in healthy plants when they start to flower (Fraser, 1981) and also during the natural aging of their leaves (Conejero *et al.*, 1979; Camacho Henriquez and Sänger, 1982b). Although all these findings show that the name PR proteins is not quite appropriate it is, nevertheless, convenient to adhere to this term as long as their nature and biological function remains unclear.

The lack of any information on the structure and structural relatedness of PR proteins prompted us to establish the complete amino acid sequence of the PR protein p14 from viroid-infected tomato plants and the partial sequence of the tobacco PR protein 1a. Because of their similarity in electrophoretic mobility, their apparent mol. wt. and their origin in related solanaceous plant species, both proteins have been assumed to be closely related. The comparison of their sequences presented here shows that this is, in fact, the case. Moreover, sequence comparison also revealed that these two PR proteins have no precedent amongst the 3061 proteins whose sequence is known at present.

Results

Sequence analysis of p14

The PR protein p14 starts to appear in the newly developing and fully expanded leaves of 'Rutgers' tomato plants 3 weeks after mechanical inoculation with potato spindle tuber viroid (PSTV) and it accumulates during the fourth week after inoculation (p.i.) to become the most abundant leaf protein in acetic acid leaf extracts (Figure 1). Therefore p14 was isolated from PSTV-infected leaves 4 weeks p.i. by acid extraction and separated from all other acid-extractable leaf proteins by ammonium sulfate precipitation and ion-exchange chromatography on sulfopropyl-Sephadex and DEAE-cellulose as previously described (Camacho Henriquez and Sänger, 1984). Direct sequencing of the purified p14 was unsuccessful because no amino acid derivatives were released from the amino terminus after Edman degradation. All attempts to deblock the N terminus of p14 with the enzyme pyroglutamic acid aminopeptidase (Podell and Abraham, 1978) failed. However, this procedure worked successfully on a cyanogen bromide fragment which contained the blocked N terminus thus identifying the blocking group as pyroglutamic acid, i.e., pyrrolidone carboxylic acid (PCA), a cyclisation product of glutamine which lacks a free α-amino group (Abraham and Podell, 1981). After removal of PCA the sequence of the first 20 amino acids could be established. For further sequence analysis, p14 was first denatured by mercaptolysis and S-carboxymethylation (SCM) at the cystine residues which were known to be present from the amino acid composition. Then fragments of SCM-p14 were generated by enzymatic digestion with different proteases and by chemical cleavage methods according to established procedures. The resulting cleavage products were isolated by gel permeation chromatography and/or reversed-phase h.p.l.c. and sequenced manually or with an automatic sequenator (Edman and Begg, 1967), as will be described in detail elsewhere. The individual fragments were then aligned according to their overlap-

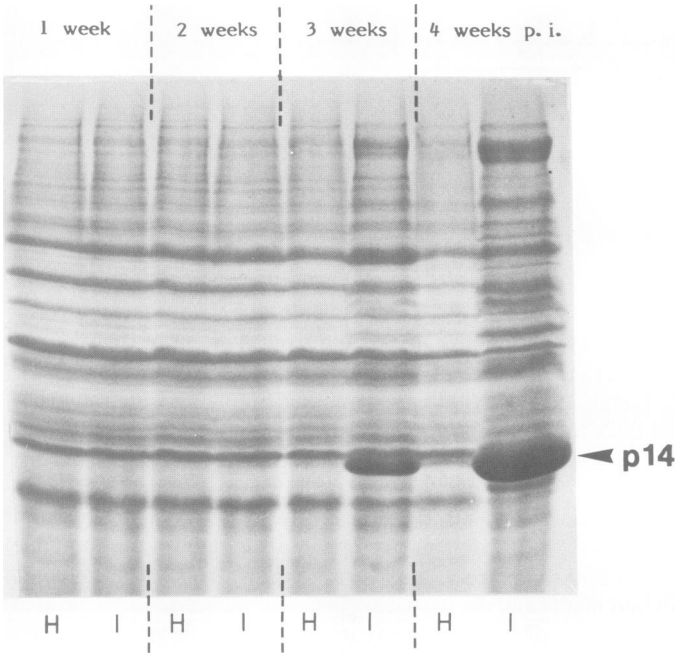

Fig. 1. Electrophoretic analysis of acid-extractable proteins from healthy **(H)** and PSTV-infected **(I)** tomato leaves demonstrating the accumulation of the pathogenesis-related (PR) protein p14 during the progress of viroid infection. Four weeks after inoculation (p.i.) p14 has accumulated to the most abundant acid-extractable leaf protein. Bands of minor PR proteins are indicated by arrows. The extraction of the proteins, their electrophoretic separation on a 20% SDS-polyacrylamide gel and their staining with Coomassie Brilliant Blue were performed as previously described (Camacho Henriquez and Sänger, 1984).

ping sequences thus yielding the complete amino acid sequence as shown in Figure 2.

Inspection of the established sequence shows that the polypeptide chain of p14 consists of 130 amino acid residues resulting in a calculated mol. wt. of 14 225, which corresponds well with the mol. wt. of 14 000 estimated from electrophoretic analysis on SDS-polyacrylamide gels (Camacho Henriquez and Sänger, 1982b, 1984). The 16 basic amino acids (10 Arg, 3 Lys, 3 His) *versus* only eight acidic amino acids (6 Asp, 2 Glu) found to be present in p14 could at least partly explain the basic nature of p14 as reflected in its isoelectric point of 10.7 (Camacho Henriquez and Sänger, 1984). The six half-cysteine residues of p14 at positions 44, 85, 91, 102, 107 and 116 are believed to participate in disulfide linkages, because iodoacetate treatment of the native non-reduced protein did not identify free sulfhydryl groups as S-carboxymethylated cysteine. When the acid hydrolysate of p14 was analysed with an amino acid analyser programmed to detect amino sugars no glucosamine or galactosamine residues were found, which are the most frequent carbohydrate components of glycosylated proteins. We therefore believe that p14 is unglycosylated, an assumption which is supported by the absence of the sequences Asn-X-Ser and Asn-X-Thr because the Asn residue in those sequences is usually N-glycosylated (Schachter, 1981). Finally, an N-terminal methionine characteristic for primary translation products is absent in p14 which indicates that it represents a processed protein.

Sequence comparison and prediction of secondary structure of p14

To establish possible sequence homologies between p14 and other proteins of known sequence, a search was carried out using the release 4.0 of the protein sequence database of February 1985 (Barker *et al.*, 1985) and the corresponding 'search' program

```
1                                                  10                                                  20
PCA-Asn-Ser-Pro-Gln-Asp-Tyr-Leu-Ala-Val-His-Asn-Asp-Ala-Arg-Ala-Gln-Val-Gly-Val-

21                                                 30                                                  40
Gly-Pro-Met-Ser-Trp-Asp-Ala-Asn-Leu-Ala-Ser-Arg-Ala-Gln-Asn-Tyr-Ala-Asn-Ser-Arg-

41                                                 50                                                  60
Ala-Gly-Asp Cys-Asn Leu-Ile-His-Ser Gly-Ala-Gly-Glu-Asn-Leu-Ala-Lys-Gly-Gly-Gly-

61                                                 70                                                  80
Asp-Phe-Thr-Gly-Arg-Ala-Ala-Val-Gln-Leu-Trp-Val-Ser-Glu-Arg-Pro-Ser-Tyr-Asn-Tyr-

81                                                 90                                                  100
Ala-Thr-Asn-Gln-Cys-Val-Gly-Gly-Lys-Lys-Cys-Arg-His-Tyr-Thr-Gln-Val-Val-Arg-Leu-

101                                                110                                                 120
Gly-Cys-Gly-Arg-Ala-Arg-Cys-Asn-Asn-Gly-Trp-Trp-Phe-Ile-Ser-Cys-Asn-Tyr-Asp-Pro-

121                                                130
Val-Gly-Asn-Trp-Ile-Gly-Gln-Arg-Pro-Tyr
```

Fig. 2. Amino acid sequence of p14 from the leaves of PSTV-infected tomato plants as established by standard methods. PCA at the N terminus denotes pyrrolidone carboxylic acid. The sequence of positions 1 – 20 was obtained by cyanogen bromide cleavage of reduced and alkylated protein at the single methionine residue Met-23 followed by enzymatic removal of the amino terminal blocking group PCA from the smaller cyanogen bromide fragment (positions 1 – 23). Positions 24 – 56 were elucidated by sequencing the larger cyanogen bromide fragment (positions 24 – 130). Positions 54 – 70 and 75 – 87 were identified by sequencing fragments obtained upon proteolytic cleavage after Glu-53 and Glu-74, respectively, with *S. aureus* protease. Positions 58 – 76 were obtained by sequencing a fragment generated by proteolytic cleavage after Lys-57 with endoprotease Lys-C. Positions 16 – 25, 93 – 99 and 107 – 121 were determined from sequence analysis of tryptic peptides and positions 86 – 94 and 98 – 110 from the analysis of peptides generated by digestion with thermolysin. Positions 120 – 130 were determined by the sequence analysis of a fragment obtained by cleavage between Asp and Pro using mild acid.

(Orcutt *et al.*, 1985). Several sequence similarities were found, but none of them exceeded a stretch of >10 residues. A closer comparison of these sequences with a computer program designed to relate sequences and to evaluate the statistical reliability of homologies (George *et al.*, 1985) failed to substantiate a relationship between p14 and any of the proteins sequenced heretofore, amongst which are a large variety of such diverse plant proteins as peroxidases, cytochromes, lectins, protease inhibitors, storage proteins, histones and virus inhibitors.

To prepare p14 for crystallographic analysis the protein was subjected to a series of different conditions under which various other proteins have been found to crystallize appropriately, but all attempts to crystallize p14 have failed so far. This failure and the lack of any proteins whose structural features could be adapted to p14 prompted us to apply the computer-aided prediction of the hydropathy and the secondary structure of this unfamiliar protein although we are fully aware of the limitations of such predictions (Nishikawa, 1983). Figure 3a clearly shows that p14 is characterized by five regions (i.e., positions 15 – 20, 42 – 48, 66 – 72, 94 – 101 and 108 – 114) in which hydrophobic amino acid residues predominate. The prediction of the secondary structure of p14 derived from four different sets of calculations as described by Rawlins *et al.* (1983) is shown in Figure 3b. From these plots it becomes evident that p14 has hydrophobic domains distributed over the molecule so that it could well interact with or span plant cell membranes.

Partial sequence of the tobacco PR protein 1a and its relationship to the sequence of p14

Amongst the PR proteins from TMV-infected tobacco leaves, three major proteins designated as PR-1a, PR-1b and PR-1c were of particular interest because the possibility existed that they are related to p14. Tobacco belongs to the same family (Solanaceae) as tomato and electrophoretic analysis had shown that the apparent mol. wts. of the three tobacco PR proteins are ~14 200 (An-

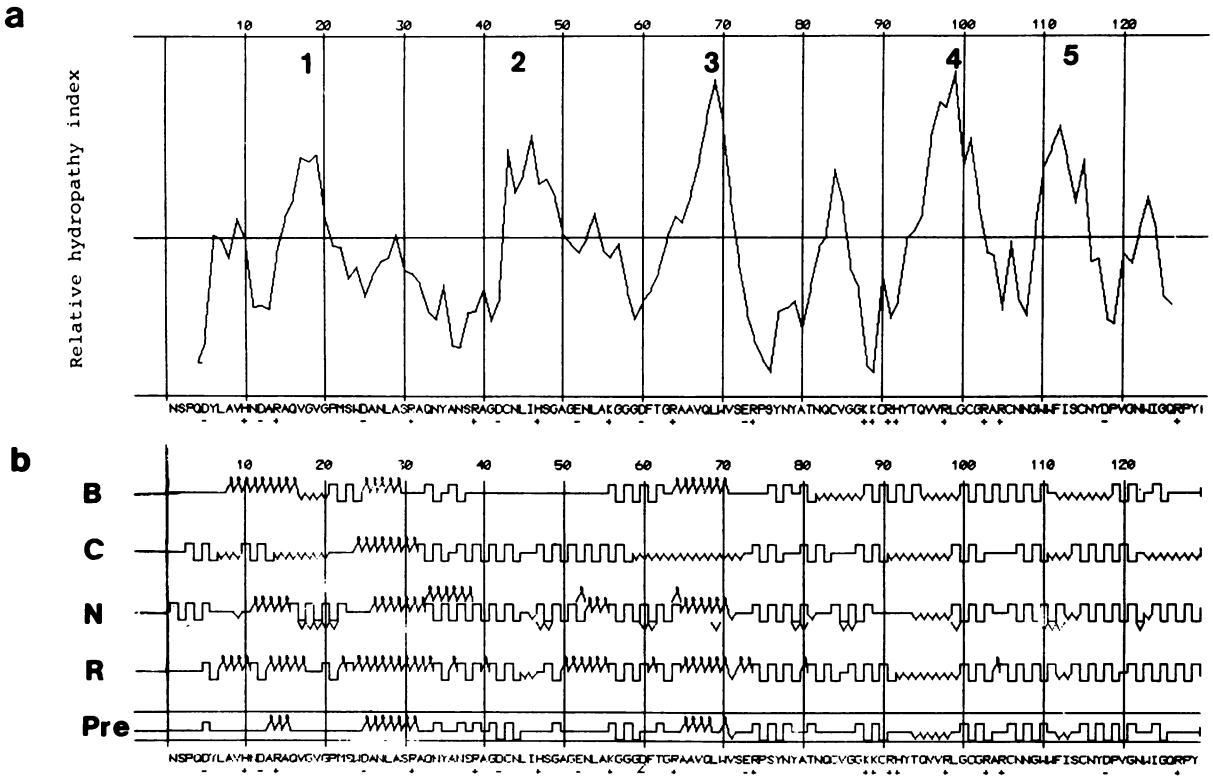

Fig. 3. Diagram plots of the hydropathy and of the joint prediction of the secondary structure of p14. **(a)** The relative hydropathy index of each residue of p14 plotted against the residue number. The indices were obtained by summing up the plotted values according to the rules of Kyte and Doolittle (1982), and the values from the contact number plot according to Nishikawa and Ooi (1980). The location of the five hydrophobic domains of the molecule are graphed above the midpoint line and indicated by numbers, whereas the hydrophilic regions are below the midpoint line. **(b)** Prediction of the secondary structure of p14 as derived from four different plots according to Rawlins *et al.* (1983). The calculations for the individual plots were made according to the rules and algorithms of the following authors: **B**, Burgess *et al.* (1974); **C**, Chou and Fasman (1978, 1979); **N**, Nagano (1977); **R**, Robson and Suzuki (1976). **PRE** shows the plot of the secondary structure of p14 as plotted when three of the four calculations gave predictions in agreement. ▮▮▮, alpha-helix; ⩙⩙, beta-sheet; ⊓⊔, beta-turn structure and −, random coil. The amino acid residues are denoted in the single letter code (IUPAC-IUB Commission, 1968).

toniw *et al.*, 1980), i.e., in the same size range as that of p14. Therefore we also started to sequence the PR-1a protein from TMV-infected tobacco leaves which had been purified and kindly provided for our analysis by Dr L.C.van Loon, Wageningen, The Netherlands.

As in the case of p14, the N terminus of the PR-1a protein was also found to be blocked by PCA. Cyanogen bromide fragments were produced and the one which contained the blocked terminus was treated with pyroglutamic acid aminopeptidase which then allowed the sequence analysis of this N-terminal fragment F1. Fragment F2, also generated by cyanogen bromide cleavage, and fragment F3, produced as a side reaction by mild acid hydrolysis, were also sequenced. The sequence derived from these three fragments comprises altogether 81 amino acid residues which corresponds to more than half of the total sequence of the PR-1a protein. The 11 acidic *versus* three basic amino acids present in the partial sequence are in accordance with the previously reported acidic nature of the PR protein 1a as reflected in its estimated isoelectric point of 4.0 (Pierpoint, 1983a). Although both PR proteins differ remarkably in charge the sequence comparison revealed (Figure 4) that the partial PR-1a protein sequences obtained could be aligned with the p14 sequence in a co-linear way without introducing any deletions or insertions and that, so far, ~60% of the sequences of both proteins are identical. This close structural relationship is further substantiated by the amino acid analysis of the S-carboxymethylated PR-1a protein which showed that, like p14, it contains six half-cysteine residues. Moreover, in the alignment the location of the two half-

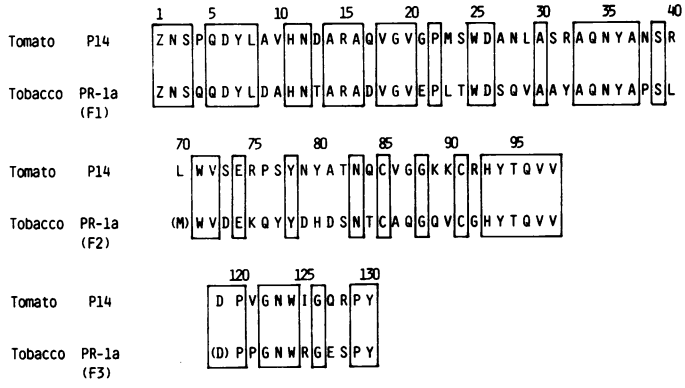

Fig. 4. Alignment of the sequences of p14 from PSTV-infected tomato leaves with PR-1a from TMV-infected tobacco leaves in the single letter code (IUPAC-IUB Commission, 1968) and using the p14 numbering as in Figure 2. In the identification of PR-1a sequences, positions 1−40 were obtained by sequence analysis of an N-terminal CNBr fragment F1 after enzymatic removal of the N-terminal blocking group pyrrolidone carboxylic acid (Z). The PR-1a sequences of positions 70−98 were determined by sequence analysis of the CNBr fragment F2. PR-1a sequence positions 119−130 were obtained from analysis of fragment F3 obtained after cleavage using mild acid. The amino acids in parenthesis were deduced by the corresponding cleavage method employed, i.e., cleavage after Met with CNBr, and between Asp and Pro using mild acid.

cysteine residues present in the sequenced fragments of the PR-1a protein corresponds to the ones in position 85 and 91 in p14 suggesting a similar disulfide bridge in the two proteins. The par-

tial PR-1a sequence also contains three of the five hydrophobic stretches where alignment with the p14 sequence would predict them. Moreover, as with p14, no amino sugars could be detected in the PR-1a protein, which indicates that it is also non-glycosylated.

From these data and from the finding that the PR proteins 1a, 1b and 1c are related to each other by amino acid composition (Antoniw *et al.*, 1980) and by serology (Matsuoko and Ohashi, 1984; Antoniw *et al.*, 1985), we conclude that p14 from tomato leaves and the three tobacco leaf proteins are members of the same family of PR proteins. This is further substantiated by recent immunological studies in which a p14 antiserum was found to cross-react with the three PR proteins from tobacco leaves (Nassuth and Sänger, 1985).

Discussion

The established sequence of p14 and its structural similarity to PR proteins from a different plant species warrants a reconsideration of the previously assumed nature and function of PR proteins in general. Since the PR proteins are unrelated to the 3061 proteins whose sequence and function are presently known, no conclusions can be drawn about the function of p14. However, the sequence of p14 clearly invalidates the assumption that the p14-related PR proteins from tobacco and other plant species are partially identical to the interferons from animal cells, as has been deduced from certain chromatographic and other physicochemical analogies (Antoniw *et al.*, 1980; Redolfi, 1983; Pierpoint, 1983a, 1983b; Abad *et al.*, 1985). Moreover, p14 and the related PR proteins neither resemble the two putative proteinaceous inhibitors of virus replication (IVR) (Loebenstein and Gera, 1981) nor the presumed anti-viral factor (AVF) from virus-infected plant tissue which is thought to be a phosphorylated glycoprotein (Sela, 1981). Moreover, p14 is not related to the disease-associated host-encoded protein with a mol. wt. of 140 000 which is found in tomato leaves after infection with four viroids and a satellite RNA (Galindo *et al.*, 1984). Finally, it is not clear whether p14 and the related PR proteins are similar to the so-called 'heat-shock proteins'. However, a dramatic accumulation of p14 as a response to supra-optimal temperatures has not been observed in tomato cell cultures (Scharf and Nover, 1982; Nover and Scharf, 1984). In fact, heat-shock proteins are rapidly synthesized upon an increase in temperature, remain inside the cell and usually disappear when adaptation has occurred (Schlesinger *et al.*, 1982). Thus they clearly differ in this respect from the PR proteins which seem to have a rather low rate of turnover (Pierpoint, 1983a) and which accumulate in the intercellular space (Parent and Asselin, 1984). Nevertheless, the accumulation of p14 and similar disease-stimulated proteins in infected, symptom-bearing and especially necrotizing leaves could well reflect a particular type of response of the host plant to the permanent stress which is imposed upon the plant by any established infection and perturbs the metabolism of the diseased foliar tissue continuously.

In conclusion, despite the wealth of phenomenological, physiological and genetic data and the different hypotheses deduced from them (see reviews by Redolfi, 1983; Antoniw and White, 1983; van Loon, 1983a, 1983b, 1985; Gianinazzi and Ahl, 1983; Pierpoint, 1983b) the function of the PR proteins, and especially their presumed involvement in acquired systemic resistance (Fraser, 1982; Fraser and Clay, 1983) are still unresolved. Since the unfamiliar sequence of a PR protein presented here has also failed to contribute to an understanding of their role, new experimental approaches are now required to unravel their presumed biological function in the diseased plant.

Materials and methods

Purification of p14

The acidic extraction of p14 from PSTV-infected tomato leaves and its subsequent purification by ammonium sulfate precipitation and ion-exchange chromatography on sulfopropyl-Sephadex and DEAE-cellulose was carried out as previously described (Camacho Henriquez and Sänger, 1984).

Analysis of the amino acid sequence

Chemical and enzymatic cleavage and modification of p14, the tobacco PR protein 1a and the subsequent sequence analysis of the overlapping fragments was carried out manually (Chang, 1981) or with an automatic sequenator (Edman and Begg, 1967) according to established methods and will be reported elsewhere in detail.

Sequence comparison

The comparison of the sequence of p14 with the published sequences of other proteins was performed with the aid of the Sequence Database Programs of the Protein Identification Resource of the National Biomedical Research Foundation, Georgetown University, Medical Center, Washington, D.C. in its site-specific computer-adapted version. In particular, Version 3.2 of the Search Program (Orcutt *et al.*, 1985), Version 2.0 of the Program for Detecting Distant Relationships (George *et al.*, 1985) and Release 4.0 of the Protein Sequence Database (Barker *et al.*, 1985) were used in combination.

Prediction of hydropathy and secondary structure

The hydropathy plot was obtained by plotting the relative hydropathy index of each residue of p14 against the residue number. The hydropathy indices were obtained by summing up the plotted values according to the rules of Kyte and Doolittle (1982) and the values from the contact number plot according to Nishikawa and Ooi (1980). The prediction of the secondary structure of p14 was obtained by the procedure of Rawlins *et al.* (1983).

Acknowledgements

We thank Dr R.Huber, Martinsried for the attempts to crystallize p14, Dr van Loon, Wageningen for providing the sample of the tobacco PR protein 1a, Mr B.Foertsch for his help in the computer-aided sequence comparison, Dr B. Wittmann-Liebold and Mr K.Ashman, Berlin for preparing the hydropathy and secondary structure plots and all our colleagues for constructive criticism and discussions.

References

Abad,P., Poupet,A., Ponchet,M., Venard,P. and Bettachini,B. (1985) *J. Chromatogr.*, **318**, 417-426.
Abraham,G.N. and Podell,D. (1981) *Mol. Cell. Biochem.*, **38**, 181-190.
Antoniw,J.F. and White,R.F. (1983) *Neth. J. Plant Pathol.*, **89**, 255-264.
Antoniw,J.F., Ritter,C.E., Pierpoint,W.S. and Van Loon,L.C. (1980) *J. Gen. Virol.*, **47**, 79-87.
Antoniw,J.F., Kueh,J.S.H., Walkey,D.G.A. and White,R.F. (1981) *Phytopathol. Z.*, **101**, 179-184.
Antoniw,J.F., White,R.F., Barbara,D.J., Jones,P. and Longley,A. (1985) *Plant Mol. Biol.*, **4**, 55-60.
Barker,W.C., Hunt,L.T., Orcutt,B.C., George,D.G., Yeh,L.S., Chem,H.R., Blomquist,M.C., Johnson,G.C., Seibel-Ross,E.I., Hong,M.K. and Ledley,R.S. (1985) *Protein Sequence Database, Release 4.0, February 15, 1985*, published by Protein Identification Resource, National Biomedical Research Foundation, Georgetown University, Medical Center, Washington, D.C.
Burgess,A.W., Ponnuswamy,P.K. and Scheraga,H.A. (1974) *Isr. J. Chem.*, **12**, 239-286.
Camacho Henriquez,A. and Sänger,H.L. (1982a) *Arch. Virol.*, **74**, 167-180.
Camacho Henriquez,A. and Sänger,H.L. (1982b) *Arch. Virol.*, **74**, 181-196.
Camacho Henriquez,A. and Sänger,H.L. (1984) *Arch. Virol.*, **81**, 263-284.
Chang,J.Y. (1981) *Biochem. J.*, **199**, 557-564.
Chou,P.Y. and Fasman,G.D. (1978) *Adv. Enzymol. Relat. Areas. Mol. Biol.*, **47**, 45-148.
Chou,P.Y. and Fasman,G.D. (1979) *Biophys. J.*, **26**, 367-384.
Conejero,V., Picazo,L. and Segado,P. (1979) *Virology*, **97**, 454-456.
Edman,P. and Begg,G. (1967) *Eur. J. Biochem.*, **1**, 80-91.
Fraser,R.S.S. (1981) *Physiol. Plant Pathol.*, **19**, 69-76.
Fraser,R.S.S. (1982) *J. Gen. Virol.*, **58**, 305-313.
Fraser,R.S.S. and Clay,C.M. (1983) *Neth. J. Plant Pathol.*, **89**, 282-292.
Galindo,J.A., Smith,D.R. and Diener,T.O. (1984) *Physiol. Plant Pathol.*, **24**, 257-275.

George,D.G., Orcutt,B.C., Dayhoff,M.C. and Barker,W.C. (1985) *Protein Sequence Database Program for Detecting Distant Relationships. Version 2.0, Jan. 28, 1985*, published by National Biomedical Research Foundation, Georgetown University, Medical Center, Washington, D.C.
Gianinazzi,S. and Ahl,P. (1983) *Neth. J. Plant Pathol.*, **89**, 275-281.
Gianinazzi,S., Martin,C. and Valée,J.C. (1970) *C.R. Hebd. Séances, Acad. Sci. Paris*, **D270**, 2383-2386.
Gianinazzi,S., Pratt,H.M., Shewry,P.R. and Miflin,B.J. (1977) *J. Gen. Virol.*, **34**, 345-351.
IUPAC-IUB Commission on Biochemical Nomenclature (1968) *Eur. J. Biochem.*, **5**, 151-153.
Kyte,J. and Doolittle,R.F. (1982) *J. Mol. Biol.*, **157**, 105-132.
Leach,J.E., Sherwood,J., Fulton,R.W. and Sequeira,L. (1983) *Physiol. Plant Pathol.*, **23**, 377-385.
Loebenstein,G. and Gera,A. (1981) *Virology*, **114**, 132-139.
Maiss,E. and Poehling,H.M. (1983) *Neth. J. Plant Pathol.*, **89**, 323.
Matsuoka,M. and Ohashi,Y. (1984) *J. Gen. Virol.*, **65**, 2209-2215.
Nagano,K. (1977) *J. Mol Biol.*, **109**, 251-274.
Nassuth,A. and Sänger,H.L. (1985) *Virus Res.*, in press.
Nishikawa,K. (1983) *Biochim. Biophys. Acta*, **748**, 285-299.
Nishikawa,K. and Ooi,T. (1980) *Int. J. Peptide Protein Res.*, **16**, 19-32.
Nover,L. and Scharf,K.-D. (1984) *Eur. J. Biochem.*, **139**, 303-313.
Orcutt,B.C., Dayhoff,M.O., George,D.G. and Barker,W.C. (1985) *Protein Sequence Database Search Program, Version 3.2, Jan. 13, 1985*, published by Protein Identification Resource, National Biomedical Research Foundation, Georgetown University, Medical Center, Washington, D.C.
Parent,J.G. and Asselin,A. (1984) *Can. J. Bot.*, **62**, 564-569.
Pierpoint,W.S. (1983a) *Phytochemistry*, **22**, 2691-2697.
Pierpoint,W.S. (1983b) *Trends Biochem. Sci.*, **8**, 5-7.
Podell,R. and Abraham,G.N. (1978) *Biochem. Biophys. Res. Commun.*, **81**, 176-185.
Rawlins,N., Ashman,K. and Wittmann-Liebold,B. (1983) *Int. J. Peptide Protein Res.*, **22**, 515-524.
Redolfi,P. (1983) *Neth. J. Plant Pathol.*, **89**, 245-254.
Robson,B. and Suzuki,E. (1976) *J. Mol. Biol.*, **107**, 327-356.
Schachter,H. (1981) in Callahan,J.W. and Lowden,J.A. (eds.), *Lysosomes and Lysosomal Storage Diseases*, Raven Press, NY, pp. 73-93.
Scharf,K.-D. and Nover,L. (1982) *Cell*, **30**, 427-437.
Schlesinger,M.J., Aliperti,G. and Kelley,P.M. (1982) *Trends Biochem. Sci.*, **7**, 222-225.
Sela,I. (1981) *Trends Biochem. Sci.*, **6**, 31-33.
van Loon,L.C. (1977) *Virology*, **80**, 417-420.
van Loon,L.C. (1983a) in Bailey,J.A. and Deverall,B.J. (eds.), *The Dynamics of Host Defense*, Academic Press, Sydney, Australia, pp. 123-190.
van Loon,L.C. (1983b) *Neth. J. Plant Pathol.*, **89**, 265-273.
van Loon,L.C. (1985) *Plant Mol. Biol.*, **4**, 111-116.
van Loon,L.C., Gianinazzi,S., White,R.F., Abu-Jawdah,Y., Ahl,P., Antoniw, J.F., Boller,T., Camacho Henriquez,A., Cohejero,V., Consirat,J.C., Goodman,R.N., Maiss,E., Redolfi,P. and Wilson,T.M.A. (1983) *Neth. J. Plant Pathol.*, **89**, 293-325.
van Loon,L.C. and van Kammen,A. (1970) *Virology*, **40**, 199-211.
Wagih,E.E. and Coutts,R.H.A. (1981) *Plant Sci. Lett.*, **21**, 61-69.

Received on 29 July 1985

Note added in proof

Our sequencing data have contributed to new experimental approaches to the PR protein problem in that they were used by Hooft van Huijsduijnen *et al.* (*EMBO J.*, **4**, 2167-2171, 1985) to construct DNA primers which allowed identification of cDNA clones of mRNAs for the PR proteins 1a, 1b and 1c in a cDNA library from TMV-infected tobacco, and to apply Northern blot analysis. It revealed that the low level of mRNAs for PR proteins already present in healthy tobacco is increased more than 100-fold upon TMV-infection, indicating that synthesis of the PR proteins is regulated at the level of transcription.