

E. González
R.I. Monsalve
X.S. Puente
M. Villalba
R. Rodríguez

Assignment of the disulfide bonds of Ole e 1, a major allergen of olive tree pollen involved in fertilization

Authors' affiliations:

E. González, R.I. Monsalve, M. Villalba and R. Rodríguez, Departamento de Bioquímica y Biología Molecular, Facultad de Química, Universidad Complutense, 28040 Madrid, Spain.

X.S. Puente, Departamento de Bioquímica, Facultad de Medicina, Universidad de Oviedo, 20006 Oviedo, Spain.

Correspondence to:

Dra Rosalía Rodríguez
Departamento de Bioquímica y Biología Molecular I
Facultad de Química
Universidad Complutense de Madrid
Madrid 28040
Spain
Fax: 34-91-3944159
E-mail: rrg@solea.quim.ucm.es

Key words: disulfide bonds; Ole e 1; olive pollen allergen; *Pichia pastoris*

Abstract: The most prevalent allergen from olive tree pollen, Ole e 1, consists of a single polymorphic polypeptide chain of 145 amino acids which includes six cysteine residues at positions 19, 22, 43, 78, 90 and 131. By using an homogeneous form of the allergen expressed in *Pichia pastoris*, the array of the disulfide bridges has been elucidated. Specific proteolysis with thermolysin and reverse-phase HPLC separation of the peptides allowed the determination of the disulfide bond between Cys43 and Cys78. Another thermolytic product, which contained three peptides linked by the remaining four cysteines, was digested with Glu-specific staphylococcal V8 protease and the products isolated by reverse-phase HPLC. Amino acid compositions and Edman degradation of the peptide products indicated the presence of the disulfide bonds at Cys19-Cys90 and Cys22-Cys131. These data can help in the analysis of the three-dimensional structure of the protein as well as in studies of its allergenic determinants.

Abbreviations: DTNB, 5,5'-dithiobis-2-nitrobenzoic acid; Ole e 1, major allergen of olive tree pollen; rOle e 1, recombinant Ole e 1; PAGE, polyacrylamide gel electrophoresis; RP, reverse phase; SDS, sodium dodecyl sulfate; S-V8 protease, Glu-specific protease from strain V8 of *Staphylococcus aureus* (EC 3.4.21.19).

Dates:

Received 25 March 1999
Revised 11 May 1999
Accepted 5 June 1999

To cite this article:

González, E., Monsalve, R.I., Puente, X.S., Villalba, M. & Rodríguez, R. Assignment of the disulfide bonds of Ole e 1, a major allergen of olive tree pollen involved in fertilization. *J. Peptide Res.*, 2000, **55**, 18–23

Copyright Munksgaard International Publishers Ltd, 2000
ISSN 1397-002X

An increasing number of individuals from developed countries are being affected by allergic diseases or IgE-mediated hypersensitivities. Diagnosis and therapy of these disorders require the use of well-characterized allergens. Plant pollens are among the most important causes of airborne allergies, being related to the geographic distribu-

tion of the vegetable source. Thus, allergy to olive tree pollen affects more than 20% of the population in areas of the Mediterranean countries (1) where the olive tree (*Olea europaea*) is widely cultivated. Although the allergen profile of olive pollen protein extracts is complex, more than 70% of the olive pollen hypersensitive patients are reactive to the major allergen Ole e 1, a polymorphic glycoprotein of 145 amino acid residues (2, 3). Its single polypeptide chain contains six cysteine residues at positions 19, 22, 43, 78, 90 and 131 (Fig. 1) which are all involved in disulfide bridges (2). The natural Ole e 1, isolated from olive tree pollen, is a highly heterogeneous protein since more than 30 positions of its polypeptide chain have been described to display micro heterogeneity (2, 3). Structure–function relationships of polymorphic proteins are difficult to resolve. Thereby, cloning and expression systems should be employed for the production of single and homogeneous molecules. Recently, we have reported on the expression of Ole e 1 in *Pichia pastoris* (4), the recombinant product being identical to the natural protein. This approach allowed us to obtain

E D V P Q P P V S Q F H I Q G Q V Y ¹⁹C ²²D T ⁴³C R A G 25
 F I T E L S E F I P G A S L R L Q ⁷⁸C K D K E N G D 50
 V T F T E V G Y T R A E G L Y S M L V E R D H K N 75
 E F ⁹⁰C E I T L I S S G R K D ¹³¹C N E I P T E G W V K 100
 P S L K F I L N T V N G T T R T V N P L G F F K K 125
 E A L P K ¹³¹C A Q V Y N K L G M Y P P N M 145

Figure 1. Amino acid sequence of Ole e 1 produced in *Pichia pastoris* and used in this study. The cysteine residues are circled.

the assignment of disulfide bridges of Ole e 1, which is a prerequisite for defining its conformation.

Besides its high structural similarity to its counterparts in other *Oleaceae* pollens such as those from privet and lilac (5, 6), Ole e 1 shows homology with pollen proteins from tomato, maize, ryegrass, *Arabidopsis thaliana*, and rice (7–9). These molecules belong to a family of pollen-specific proteins, which have been suggested to play a role in pollen hydration and/or pollen germination (10). This family of proteins has also been proposed to be related to the soybean trypsin

Table 1. Amino acid composition of cystine-containing peptides from the thermolytic (Th peptides) and S-V8 protease (SV peptides) treatments of rOle e 1

	Th10	Th12	Th15	SV4	SV5	SV6	SV8
Cys ^a	2.2 (2)	2.1 (2)	3.9 (4)	1.5 (2)	2.0 (2)	2.0 (2)	2.0 (2)
Asx	3.5 (3)	3.1 (3)	3.1 (3)	2.5 (3)		2.9 (3)	1.5 (1)
Thr		1.0 (1)	1.8 (2)		1.0 (1)		
Ser			1.0 (1)			1.8 (2)	
Glx	3.4 (3)	3.3 (3)	1.8 (2)	1.2 (1)		1.1 (1)	2.9 (3)
Pro			1.4 (2)		1.0 (1)		
Gly	1.3 (1)	2.1 (2)	3.0 (3)	1.4 (1)	1.2 (1)	1.5 (1)	
Ala			1.8 (2)		2.0 (2)		
Val		1.0 (1)	0.9 (1)	0.8 (1)		1.1 (1)	
Met							
Ile			1.3 (2)		1.0 (1)	0.9 (1)	
Leu	2.4 (2)	1.0 (1)	1.0 (1)				1.0 (1)
Tyr			- (1 ^b)	- (1 ^b)		- (1 ^b)	
Phe	1.0 (1)	1.0 (1)					0.8 (1)
His							
Lys	2.0 (2)	1.8 (2)	1.8 (2)	1.0 (1)	1.0 (1)	1.0 (1)	1.9 (2)
Arg	1.0 (1)		3.4 (2)	1.0 (1)	1.0 (1)	1.1 (1)	
Trp			- (1 ^b)				
Sequence	39–50	41–52	17–25	17–20	21–25	17–20	41–47
position	77–79	77–79	83–98	86–92	127–131	83–92	77–79
			127–131				

a. Determined as cysteic acid after performic acid oxidation.

b. Determined spectrophotometrically.

inhibitor (Kunitz) family (9, 10), although their members lack the active site for catalysis (9). Since the disulfide arrangement is a typical feature for the classification of the protease inhibitor protein family (11), the elucidation of those of Ole e 1 might clarify this putative homology.

Experimental Procedures

Ole e 1 from pollen was isolated as described previously (3). Recombinant Ole e 1 (rOle e 1) was expressed in *P. pastoris* as described (4). rOle e 1 was purified from the extracellular medium on a DEAE-cellulose column and then subjected to size exclusion chromatography on a Sephadex G-75 column. Characterization of rOle e 1 was achieved as described previously (4): circular dichroism and fluorescence spectra, amino acid composition and sequence, mass spectra, and immunologic analysis were carried out, in comparison with those of the natural Ole e 1, in order to assess the correct structure and folding of the expressed protein.

Peptides from rOle e 1 containing disulfide bonds were isolated by a two-step procedure of proteolytic treatments with thermolysin and Glu-specific staphylococcal V8 (S-V8) proteases. This procedure was also followed with the natural allergen, to compare the data obtained with the recombinant protein.

Thermolysin digestion was performed according to a previous report (12) with minor modifications. Purified rOle e 1 (200 µg) was dissolved in 0.1 M ammonium acetate buffer, pH 6.5, containing 10 mM CaCl₂. Thermolysin was added to this solution, at an enzyme/substrate ratio of 1 : 20 (by mass). Digestion was carried out for 16 h at 37°C. The mixture of peptides was separated by RP-HPLC on a Nucleosil C-18 column with an acetonitrile gradient in 0.1% trifluoroacetic acid at 1 mL/min flow rate. The elution was monitored by measuring absorption at 214 nm wavelength.

The thermolytic digestion was done again, but on a larger scale (from 1 mg of rOle e 1), in order to isolate the peptide Th₁₅, which contained four half cystinyl residues. This peptide was purified on a Sephadex G-25 column (1 × 29 cm), equilibrated in 0.2 M ammonium bicarbonate, pH 8.0. The eluate was monitored at 214 nm. As peptide Th₁₅ contains tryptophan, it was also detected at 280 nm and analyzed by SDS-PAGE. Afterwards, Th₁₅ (150 µg) was digested with Glu-specific S-V8 protease in 25 µL of 0.1 M sodium phosphate buffer, pH 7.8, containing 2 mM EDTA, at 37°C for 16 h (13). The digested sample was separated on a RP-HPLC C-18 column under conditions identical to those described above for thermolytic peptides.

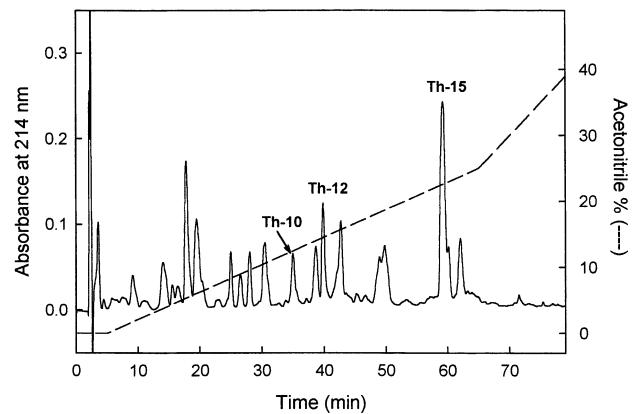


Figure 2. RP-HPLC elution profile of the digestion of Ole e 1 with thermolysin, by using a gradient of acetonitrile in 0.1% trifluoroacetic acid. Peaks showing absorbance at 214 nm were manually collected and analyzed. The cystine-containing peptides are numbered.

Amino acid compositions were determined by automatic analysis of the hydrolyzed samples by standard procedures. Peptides (1–5 nmol) were hydrolyzed at 105°C for 24 h, with 5.7 N HCl containing 0.1% phenol and 20 nmol/mL 2-amino hexanoic acid as internal standard, in sealed tubes under vacuum. Hydrolyzed samples were analyzed on a Beckman system 6300 amino acid analyzer with an analog interface module of System Gold. The cysteine content was determined as cysteic acid after performic acid oxidation. Tryptophan and tyrosine were spectrophotometrically detected in the peptides before acid hydrolysis. Free sulfhydryl groups were determined after hydrolysis as carboxymethyl-cysteine from the alkylated protein performed under non-reducing conditions, as well as by reaction with 5,5'-dithiobis-2-nitrobenzoic acid (DTNB), in 0.1 M Tris, pH 7.5, containing 1 mM EDTA and 2% SDS (14). The molar absorption coefficient at 214 nm was employed for calculations.

NH₂-terminal Edman degradations were automatically performed by using an Applied Biosystems model 477A sequencer. The resultant phenylthiohydantoin derivatives were analyzed by HPLC with the standard Applied Biosystems program (15).

Analytical gel electrophoresis was performed according to Laemmli (16) in 15% polyacrylamide gels. Treatment of the samples with 2-mercaptoethanol, for 20 min at 80°C, was alternatively performed. Proteins were stained with Coomassie blue reagent.

Results

rOle e 1, expressed in *P. pastoris*, is a soluble and properly folded protein whose structural and immunologic properties

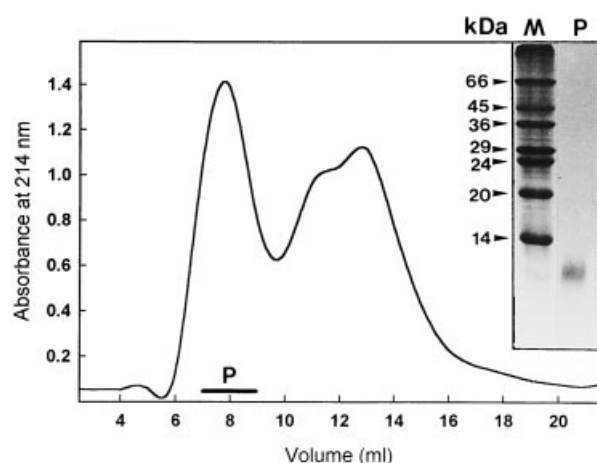


Figure 3. Preparative chromatographic purification of peptide Th15 on a Sephadex G-25 column. Fractions indicated by the horizontal bar (P) were collected and lyophilized. Trp content was spectrophotometrically assessed. Inset: (P) SDS-PAGE of sample P; (M) molecular mass markers.

were identical to that of the natural allergen isolated from pollen (4). The analysis of free sulfhydryl groups in rOle e 1 with both DTNB and iodoacetamide indicated that all the cysteine residues must be involved in disulfide bonds, as in the natural allergen (3). The recombinant allergen was digested with thermolysin and the peptides separated in a RP-HPLC column. The elution profile is shown in Fig. 2. Thermolytic proteolysis rendered 16 well-defined peaks. Three of them (Th10, Th12 and Th15) were found to be the cystine-containing peptides. Table 1 shows the amino acid composition of these peptides. On the basis of the amino acid sequence of Ole e 1 (Fig. 1), it can be inferred that the amino acid compositions of peptides Th10 and Th12 correspond to the same region of the protein, and contain the two cysteines at positions 43 and 78 of the polypeptide chain. The amino acid composition of Th15 was in agreement with the overall content of three peptides: Val17 to Gly25, Ile83 to Trp98, and Ala127 to Cys131. Therefore, Th15 should contain the four remaining cysteine residues, located at positions 19, 22, 90 and 131. This peptide also exhibited a significant absorbance at 280 nm, confirming that it contained the single Trp residue of the protein.

The data from automatic Edman degradation of Th12, which are shown in Table 2, confirmed the disulfide arrangement between the two cysteine residues (Cys43 and Cys78) contained in the segments Leu41 to Thr52 and Phe77 to Glu79. This disulfide bridge was the only one clearly assigned in complementary studies with the natural Ole e 1 isolated from pollen. Subsequent digestions of this protein rendered very complex chromatographic profiles (data not shown), due to its inherent polymorphism.

Table 2. Amino acid sequence of cystine-containing peptides. Underlined letters indicate residues determined by Edman degradation

Peptides	Sequence	Position in sequence
Th12	<u>LQCKDK</u> ENG DVT	41–52
	<u>FCE</u>	77–79
SV4	<u>VYCD</u>	17–20
	<u>GRKDCNE</u>	86–92
SV5	<u>TCRAG</u>	21–25
	<u>ALPKC</u>	127–131
SV8	<u>LQCKDKE</u>	41–47
	<u>FCE</u>	77–79

In order to obtain a sufficient amount of peptide Th15, 1 mg of rOle e 1 was digested with thermolysin and chromatographed on a Sephadex G-25 column. The elution profile is shown in Fig. 3. Fractions containing Th15 were detected at 280 nm and analyzed by SDS-PAGE. The molecular mass of this peptide, calculated on the basis of its amino acid composition, is 3450 Da, which agrees approximately with its electrophoretic mobility on the polyacrylamide gel (Fig. 3, inset). This band was not observed when the sample was previously treated with 2-mercaptoethanol. Fractions containing Th15 were pooled, lyophilized and digested with S-V8 protease under conditions for cleavage at both Glu and Asp residues (13). The digestion products were separated by RP-HPLC, as shown in Fig. 4, these being SV4, SV5, SV6 and SV8 the cystine-containing peptides. Their amino acid compositions are shown in Table 1. From these results, it can be deduced that SV4 and SV6 came from the same segment of the protein, and could correspond to peptides Val17 to Asp20 plus Gly86 to Glu92, and Val17 to Asp20 plus Ile83 to Glu92, respectively. This result indicates the existence of partial hydrolysis of the peptide bonds sensitive to the protease, and that both peptides, SV4 and SV6, must contain the disulfide bond between Cys19 and Cys90. Edman degradation of SV4 (Table 2) confirmed the arrangement of this disulfide bridge.

Peptide SV5 exhibits an amino acid composition consistent with that of peptides Thr21 to Gly25 plus Ala127 to Cys131. This correspondence was confirmed by Edman degradation (Table 2) indicating that SV5 contains the disulfide bond between Cys22 and Cys131. The amino acid composition and Edman degradation data of the peak SV8, which correspond to peptides Leu41 to Glu47 plus Phe77 to Glu79, confirmed the data obtained for peptide Th12, that contains the disulfide bridge linking Cys43 and Cys78.

Taking into account the above results, the disulfide structure of Ole e 1 can be represented as shown in Fig. 5.

Discussion

Disulfide bonds play an important role in the process of protein folding and are of particular importance for tertiary structure stabilization (17, 18). In addition, the disulfide bonding pattern can provide useful information for directing mutagenesis and fragmentation studies aimed at defining functional domains of proteins (19, 20). In this work, we report on the determination of the disulfide bridges of Ole e 1, a major allergen from olive pollen, using a recombinant form obtained in *Pichia pastoris* (4).

Ole e 1 is a highly polymorphic glycoprotein (2, 3). RP-HPLC peptide maps of natural Ole e 1 show a number of peptides remarkably higher than that expected for an homogeneous protein, most of them appearing at low yield. Besides polymorphism, two additional difficulties hindered the complete elucidation of the disulfide bonds of the natural protein: the stability of the native structure of Ole e 1 against specific proteases, and the proximity of two half cystines in the sequence Cys19-Asp-Thr-Cys22 (Fig. 1).

In order to obtain a unique molecular species of Ole e 1, a cDNA encoding the allergen was expressed in *E. coli* as a fusion protein with a connecting peptide sensitive to thrombin (2). However, only 1% of the product was immunologically active, most of the protein being an aggregate of high molecular weight. Recently, a soluble and properly folded recombinant form of Ole e 1 has been

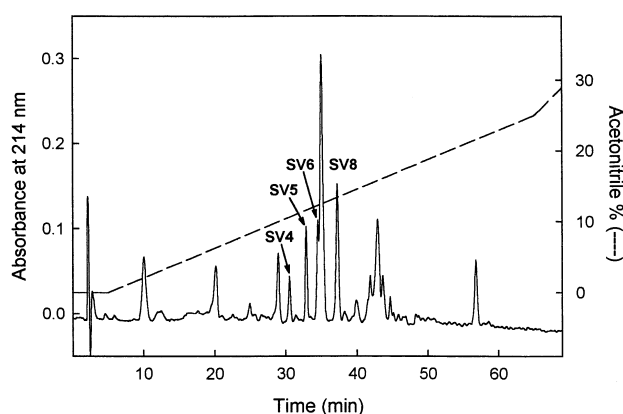


Figure 4. RP-HPLC elution profile of the S-V8 protease digestion of peptide Th15, by using a gradient of acetonitrile in 0.1% trifluoroacetic acid. Ten peaks were manually collected for amino acid analysis and sequencing. The cystine-containing peptides are numbered as SV4, SV5, SV6 and SV8.

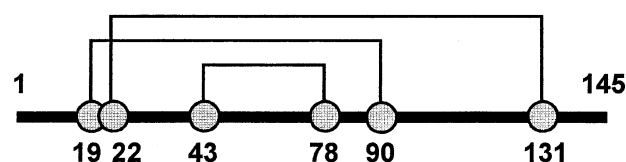


Figure 5. Topological representation of the disulfide bonds of Ole e 1. A solid circle indicates a half cystine residue at that position.

produced in *P. pastoris* (4). The non-polymorphic nature of the recombinant protein allowed the elucidation of the disulfide linkages by combining exhaustive protease treatments and RP-HPLC separations. By these means, clear peptides at convenient yields were obtained. The three disulfide bonds are established between Cys19-Cys90, Cys22-Cys131, and Cys43-Cys78 (Fig. 5), which can also be referred to as: 1-5, 2-6 and 3-4, considering the relative positions of the cysteine residues in the polypeptide chain.

Ole e 1 belongs to a family of pollen proteins that have been demonstrated to be involved in fertilization processes (10). Although identity among phylogenetically unrelated members of this family of proteins, such as maize, tomato and olive tree, is around 35%, their alignment involves identical positions for the six cysteine residues (3, 9). Therefore, it can be assumed that all members of these pollen proteins would display identical disulfide arrangement.

In addition, homology between this family of pollen proteins and the soybean trypsin inhibitor (Kunitz) family (9, 10) has also been suggested. The Kunitz family exhibits a characteristic disulfide bond arrangement (11), between the cysteine residues 1-2 and 3-4 (using the schematic nomenclature mentioned above). This arrangement differs from that determined for Ole e 1. Thus, and taking together the absence of critical residues at the active site (9), we suggest that the proposed homology, between Kunitz protein family and Ole e 1-related pollen proteins, should be reconsidered.

The elucidation of the disulfide array of Ole e 1 will help in future studies, such as the determination of its three-dimensional structure and defining its B- and T-cell epitopes (21, 22). It will also be useful for designing recombinant DNA strategies to produce engineered allergens to be used in safer and more effective immunotherapy.

Acknowledgments: This work has been supported by the Dirección General de Investigación Científica y Técnica (Spain) (PM95-0074). E. G. is the recipient of a pre-doctoral fellowship from the Ministerio de Educación y Cultura (Spain).

References

1. Bousquet, J., Cour, P., Guerin, B. & Michel, F.-B. (1984) Allergy in the Mediterranean area. I. Pollen counts and pollinosis of Montpellier. *Clin. Allergy* **14**, 249–258.
2. Villalba, M., Batanero, E., Monsalve, R.I., González de la Peña, M.A., Lahoz, C. & Rodríguez, R. (1994) Cloning and expression of Ole e I, the major allergen from olive tree pollen. Polymorphism analysis and tissue specificity. *J. Biol. Chem.* **269**, 15217–15222.
3. Villalba, M., Batanero, E., López-Otín, C., *et al.* (1993) The amino acid sequence of Ole e I, the major allergen from olive tree (*Olea europaea*) pollen. *Eur. J. Biochem.* **216**, 863–869.
4. Huecas, S., Villalba, M., González, E., Martínez-Ruiz, A. & Rodríguez, R. (1999) Production and detailed characterization of biologically active olive pollen allergen Ole e 1 secreted by the yeast *Pichia pastoris*. *Eur. J. Biochem.* **261**, 539–545.
5. Batanero, E., González de la Peña, M.A., Villalba, M., Monsalve, R.I., Martín-Esteban, M. & Rodríguez, R. (1996) Isolation, cDNA cloning and expression of Lig v 1, the major allergen from privet pollen. *Clin. Exp. Allergy* **26**, 1401–1410.
6. Batanero, E., Villalba, M., López-Otín, C. & Rodríguez, R. (1994) Isolation and characterization of an olive allergen-like protein from lilac pollen. Sequence analysis of three cDNA encoding protein isoforms. *Eur. J. Biochem.* **221**, 187–193.
7. Twell, D., Wing, R., Yamaguchi, J. & McCormick, S. (1989) Isolation and expression of an anther-specific gene from tomato. *Mol. Gen. Genet.* **217**, 240–245.
8. Hanson, D.D., Hamilton, D.A., Travis, J.L., Bashe, D.M. & Mascarenhas, J.P. (1989) Characterization of a pollen-specific cDNA clone from *Zea mays* and its expression. *Plant Cell* **1**, 173–179.
9. van Ree, R., Hoffman, D.R., van Dijk, W., *et al.* (1995) Lol p XI, a new major grass pollen allergen, is a member of a family of soybean trypsin inhibitor-related proteins. *J. Allergy Clin. Immunol.* **95**, 970–978.
10. Muschietti, J., Dircks, L., Vancanneyt, G. & McCormick, S. (1994) LAT52 protein is essential for tomato pollen development: pollen expressing antisense LAT52 RNA hydrates and germinates abnormally and cannot achieve fertilization. *Plant J.* **6**, 321–338.
11. Laskowski Jr, M. & Kato, I. (1980) Protein inhibitors of proteinases. *Ann. Rev. Biochem.* **49**, 593–626.
12. Heinrikson, R.L. (1977) Applications of thermolysin in protein structural analysis. *Methods Enzymol.* **47**, 175–189.
13. Drapeau, G.R. (1976) Protease from *Staphylococcus aureus*. *Methods Enzymol.* **45**, 469–475.
14. Ellman, G.L. (1959) Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* **82**, 70–77.
15. Hewick, R.M., Hunkapiller, M.W., Hood, L.E. & Dreyer, W. (1981) A gas-liquid solid phase peptide and protein sequenator. *J. Biol. Chem.* **256**, 7990–7997.
16. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
17. Creighton, T.E. (1988) Disulphide bonds and protein stability. *Bioessays* **8**, 57–63.
18. Gehrmann, J., Alewood, P.F. & Craik, D.J. (1998) Structure determination of the three disulfide bond isomers of alpha-conotoxin GI: a model for the role of disulfide bonds in structural stability. *J. Mol. Biol.* **278**, 401–415.
19. Leonard, C.K., Spellman, M.W., Riddle, L., Harris, R.J., Thomas, J.N. & Gregory, T.J. (1990) Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *J. Biol. Chem.* **265**, 10373–10382.
20. Ikegaya, K., Hirose, M., Ohmura, T. & Nokihara, K. (1997) Complete determination of disulfide forms of purified recombinant human serum albumin, secreted by the yeast *Pichia pastoris*. *Anal. Chem.* **69**, 1986–1991.
21. Nishiyama, C., Yuuki, T., Takai, T., Okumura, Y. & Okudaira, H. (1993) Determination of three disulfide bonds in a major house dust mite allergen, Der f II. *Int. Arch. Allergy Immunol.* **101**, 159–166.
22. Takai, T., Yokota, T., Yasue, M. *et al.* (1997) Engineering of the major house dust mite allergen Der f 2 for allergen-specific immunotherapy. *Nat. Biotechnol.* **15**, 754–758.