

NMR Solution Structure of Ole e 6, a Major Allergen from Olive Tree Pollen*

Received for publication, June 1, 2004, and in revised form, June 30, 2004
Published, JBC Papers in Press, July 7, 2004, DOI 10.1074/jbc.M406045200

Miguel Ángel Treviño‡§, María Flor García-Mayoral‡§, Patricia Barral¶||, Mayte Villalba¶, Jorge Santoro‡, Manuel Rico‡, Rosalía Rodríguez¶||, and Marta Bruix‡**

From the ‡Departamento de Espectroscopía y Estructura Molecular, Instituto de Química Física “Rocasolano,” Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain and the ¶Departamento de Bioquímica y Biología Molecular I, Facultad de Química, Universidad Complutense, 28040 Madrid, Spain

Ole e 6 is a pollen protein from the olive tree (*Olea europaea*) that exhibits allergenic activity with a high prevalence among olive-allergic individuals. The three-dimensional structure of Ole e 6 has been determined in solution by NMR methods. This is the first experimentally determined structure of an olive tree pollen allergen. The structure of this 50-residue protein is based on 486 upper limit distance constraints derived from nuclear Overhauser effects and 24 ϕ torsion angle restraints. The global fold of Ole e 6 consists of two nearly antiparallel α -helices, spanning residues 3–19 and 23–33, that are connected by a short loop and followed by a long, unstructured C-terminal tail. Viewed edge-on, the structured N terminus has a dumbbell-like shape with the two helices on the outside and with the hydrophobic core, mainly composed of 3 aromatic and 6 cysteine residues, on the inside. All the aromatic rings lie on top of and pack against the three disulfide bonds. The lack of thermal unfolding, even at 85 °C, indicates a high conformational stability. Based on the analysis of the molecular surface, we propose five plausible epitopes for IgE recognition. The results presented here provide the structural foundation for future experiments to verify the antigenicity of the proposed epitopes, as well as to design novel hypoallergenic forms of the protein suitable for diagnosis and treatment of type-I allergies. In addition, three-dimensional structure features of Ole e 6 are discussed to provide a basis for future functional studies.

Type-I allergies are one of the most widespread pathologies in developed countries where they affect more than 20% of the population, provoking respiratory, skin, and intestinal disorders. The key factor responsible for the disease is the release of inflammatory mediators, such as histamine (1–3), after an IgE-mediated immunological response (4). The diagnosis and

treatment of allergies are based on the use of extracts and vaccines, which require large amounts of biological materials obtained from natural sources for their production. These extracts can provoke lateral effects such as anaphylactic shock, are expensive to obtain, and do not assure a constant composition of allergens in extracts. Therefore, the production of the allergens via peptidic synthesis or molecular biology techniques could be a more effective and cheaper alternative source of allergens that could permit the standardization of treatments (5). The identification of the epitopes, regions of the antigen recognized by the immunoglobulin, is essential in order to design these synthetic vaccines. Until now, epitope-mapping techniques have been based on the search for linear or sequential epitopes (6–9), but they have failed in the identification of discontinuous or conformational epitopes. The shape and accessibility of each region of the allergen are essential for recognition by the IgEs (10), so to determine the structure of the allergens is a necessary step in order to define accurately the putative epitopes. This way we can limit the putative residues preferentially recognized by the IgEs, minimize the number of epitopes to be tested for mutating them, and verify their implication in the antigenic response to prepare hypoallergenic forms of the protein to be used as vaccines.

Pollinosis, type-I allergies caused by anemophilous pollens, are one of the most widespread allergies (11), and olive tree (*Olea europaea*) pollen is one of the main causes of seasonal respiratory allergy in Mediterranean countries (12). Important cross-reactions with other Oleaceae pollens, like ash (*Fraxinus excelsior*), lilac (*Syringa vulgaris*), or privet (*Ligustrum vulgare*), have been described previously (13–19). At the moment, 10 allergens from olive pollen, Ole e 1–10, have been described (20–24), but their three-dimensional structures have not been determined. Ole e 6 is a 50-amino acid protein composed of a unique polypeptide chain with unknown function. Its allergenic prevalence varies for different geographical areas and reaches 55% in areas where the olive tree is extensively cultivated (21, 25). The amino acid sequence contains two sets of the sequential motif Cys-X₃-Cys-X₃-Cys (25) and shows similarity with a polypeptide sequence deduced from tobacco (*Nicotiana tabacum*) cDNA. This Cys-enriched motif has also been found in a putative protein deduced from the *Caenorhabditis elegans* genomic structure. Ole e 6 has not been shown to share IgE epitopes with other allergenic pollens, which makes Ole e 6 a good candidate to be included in synthetic vaccines for olive pollinosis because there are not expected significant homologues or similar sequences in other plants.

Recently, we have produced in *Pichia pastoris* and characterized the recombinant form of Ole e 6 (22). In this work we have used its ¹⁵N-labeled form to determine its high resolution three-dimensional structure in solution. The analysis of this

* This work was supported by Grant CAM2002-07B/0054 from the Comunidad Autónoma de Madrid and Grant SAF2002-02711 from the Ministerio de Ciencia y Tecnología (Spain). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1SS3) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

§ Recipients of fellowships from the Comunidad Autónoma de Madrid (CAM, Spain).

|| Recipient of a fellowship from the Ministerio de Educación, Cultura y Deporte (Spain).

** To whom correspondence should be addressed. Fax: 34-91-561-94-00; E-mail: mbruix@iqfr.csic.es.

structure is essential to infer information about the accessibility, shape, and mobility of the different regions of the molecule, which can lead to the design of new vaccines specific for olive tree pollen. In addition, the structure determination can also provide a basis to understand or hypothesize on the biological function of this novel plant protein.

EXPERIMENTAL PROCEDURES

Protein Sample—The cDNA from Ole e 6 was obtained as described previously (25), and it was cloned into the vector pPIC9 to allow its expression in *P. pastoris*. The plasmid containing the cDNA was integrated in the *P. pastoris* GS115 his4 strain genome, and the production of the recombinant Ole e 6 (rOle e 6)¹ was carried out as described previously (26). Briefly, a colony of plasmid-containing *P. pastoris* was grown in 10 ml of BMG (100 mM K₂HPO₄, pH 6, 0.34% yeast nitrogen base, 1% (NH₄)₂SO₄, 4 × 10⁻⁵% biotin, and 1% glycerol) for 2 days at 30 °C. This culture was used as the preinoculum of a 1-liter culture in the same medium and conditions. After 2 days, the cells were grown in 200 ml of BMM, which was the same as BMG but with 0.5% methanol instead of glycerol, to induce the production of the recombinant protein. After a 4-day culture, the supernatant was collected and dialyzed in the presence of 20 mM NH₄HCO₃. The dialyzed protein was separated by size through a Sephadex G-50 chromatography column and later by reverse phase HPLC through a nucleosyl-C₁₈ column.

Collected peaks from HPLC were resolved in 17% SDS-PAGE, and after Western blotting the peak containing the pure protein was determined. The protein from this peak was selected and used in the subsequent experiments. ¹⁵N-rOle e 6 was produced and isolated in the same way but substituting (NH₄)₂SO₄ with (¹⁵NH₄)₂SO₄ in the BMG and BMM media. All the samples were analyzed by amino acid analysis, N-terminal sequencing by Edman degradation, and mass spectroscopy.

Circular Dichroism Analyses—CD spectra of rOle e 6 were obtained on a Jasco J-715 spectropolarimeter (Japan Spectroscopic Co., Tokyo, Japan), fitted with a 150-W xenon lamp, and connected to a Neslab RTE-111 thermostabilizer bath. Far-UV spectra were registered in the range of 190–250 nm using an optical cell path of 0.1 cm. The spectra were recorded at 50 nm/min scanning speed, and the accumulation of four spectra was achieved. Samples were analyzed in 20 mM sodium phosphate, pH 7.2, at 20 and 85 °C, and the protein concentration was 0.2 mg/ml. Mean residue mass ellipticities were calculated based on 116 as the average molecular mass/residue, obtained from the amino acid composition of Ole e 6, and expressed in terms of θ (degrees·cm²·dmol⁻¹). Final spectra were corrected by subtracting the corresponding base-line spectrum obtained for the buffer under identical conditions.

Temperature-induced denaturation of rOle e 6 was monitored by CD measurements at 220 nm. The temperature was increased from 20 to 85 °C and then cooled to 20 °C at 0.5 °C/min.

NMR Spectroscopy—rOle e 6 samples were prepared for NMR experiments at 0.5–1 mM in 95% H₂O, 5% D₂O or in D₂O containing sodium-4,4-dimethyl-4-silapentane-1-sulfonate at pH 6.0. NMR spectra were carried out at 278, 298, or 308 K on a Bruker AV600 NMR spectrometer equipped with a triple-resonance cryoprobe and an active shielded z-gradient coil. Chemical shifts were referenced to the internal sodium-4,4-dimethyl-4-silapentane-1-sulfonate as described previously (27). On the unlabeled sample two-dimensional COSY (28), total correlation spectroscopy (65-ms mixing time) (29), and NOE spectroscopy (150-ms mixing time) (30) spectra were acquired in H₂O and D₂O. Two-dimensional ¹⁵N HSQC (31) and three-dimensional nuclear Overhauser effect spectroscopy-HSQC (80-ms mixing time) (32) spectra were acquired in H₂O on the ¹⁵N sample. Additionally, ¹⁵N HSQC spectrum was recorded without decoupling during acquisition, using an α/β half-filter to detect the low field and high field components of the multiplets (33). To test the association state of the sample, one-dimensional ¹H spectra were run at different sample concentrations from 1 to 0.02 mM. No changes in chemical shifts or line widths were observed upon dilution, so we conclude that the sample was monomeric under the conditions described above. The processing of the spectra was carried out by using the program XWIN-NMR (Bruker Biospin, Karlsruhe, Germany) and NMRPipe (34). The spectra analysis, assignment, and cross-peak volume calculations were performed with ANSIG 3.3 (35).

¹ The abbreviations used are: rOle e 6, recombinant Ole e 6; HPLC, high pressure liquid chromatography; W, watt; HSQC, heteronuclear single quantum correlation spectroscopy; NOE, nuclear Overhauser effect.

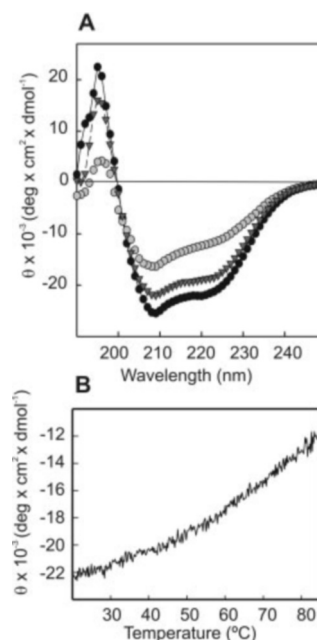


FIG. 1. Temperature-induced changes in the secondary structure of rOle e 6. A, far-UV CD spectra of rOle e 6 at 20 °C (black circles), 85 °C (gray circles), and after cooling to 20 °C (inverted gray triangles). B, residue mass ellipticity measurements at 220 nm during the increasing of the temperature from 20 to 85 °C.

Structure Calculation—1428 unambiguous NOEs were converted into 900 upper limits by the CALIBA program (36). The disulfide bond pattern of Ole e 6 was unambiguously determined on the basis of NOE evidences. For the three S–S pairs (8–34, 12–30, and 16–26), at least one intercyysteine NOE could be found, and the corresponding restraints were introduced in the calculation. After DYANA filtering, 486 conformationally relevant distances were obtained. These constraints were supplemented with 24 ϕ angle restraints (–200° to –10°) derived from the cross-correlated dipole-dipole relaxation between ¹HN-¹⁵N and ¹HN-¹H α (33). The negative ϕ angles correspond to the following residues: Ala-3, Gln-4, Phe-5, Lys-19, Asn-21, Thr-24, Cys-26, Cys-30, Thr-32, Cys-34, Ser-35, Val-36, Lys-37, Asp-38, Val-39, Lys-40, Glu-41, Lys-42, Leu-43, Glu-44, Tyr-46, Lys-47, Lys-49, and Asn-50. In the final steps, stereospecific limits were introduced using the program GLOMSA (36) for the H β of residues Phe-5, Tyr-9, His-13, Cys-16, Asp-18, Phe-25, Cys-26, Lys-29, Cys-30, and Asp-31 and for H α of Gly-20.

The three-dimensional structure of rOle e 6 was determined with the program DYANA 1.5 (37). A total of 50 structures were calculated, and the 25 structures with the lowest target function were refined by restrained energy minimization in vacuum (500 steps) using the program AMBER7 (38). The program MOLMOL (39) was used to visualize and globally characterize the structures. The quality of the refined structures was evaluated with the program PROCHECK-NMR (40).

RESULTS

Thermal Stability—To test the thermal stability of rOle e 6, CD spectra were obtained at 20 °C, after heating at 85 °C, and again at 20 °C after cooling (Fig. 1A). Temperature-induced changes in the secondary structure of rOle e 6 were registered by measure of CD values at 220 nm between 20 and 85 °C (Fig. 1B). The experiment showed a linear increase in the mean residue mass ellipticity without a transition between two states, making impossible the determination of a T_m value. The CD spectra obtained at 85 °C showed that the protein maintained elements of secondary structure. The recovery of the initial temperature induced almost the complete recuperation of the original spectrum, indicating the restoration of the structure of the allergen.

Description of the Ole e 6 Structure—The assignments of the ¹H and ¹⁵N resonances were obtained following the standard strategy (41) and have been deposited in the BioMagResBank

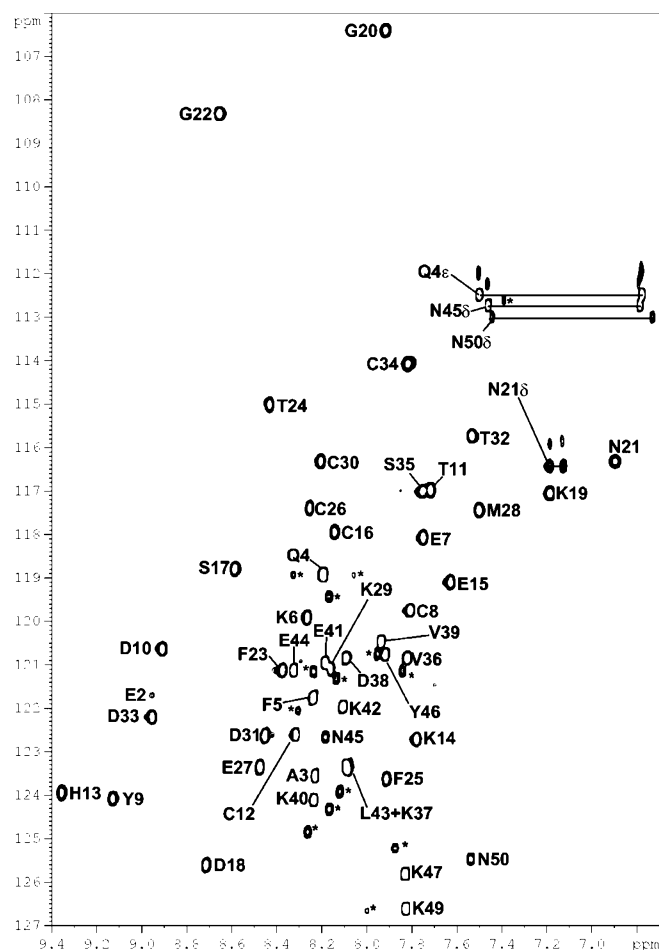


FIG. 2. 600 MHz ^1H - ^{15}N HSQC spectrum of rOle e 6 (pH 6.0, 35 °C). All correlations have been assigned and labeled according to residue type and sequence number, and the atom type is also indicated for side chain correlations (those from Asn and Gln joined by horizontal lines). Non-identified peaks are labeled with an asterisk.

data base (42) under the accession number BMRB-6139. The ^1H and ^{15}N assignments for the backbone and side chains are complete, except for Asp-1. As an example, the assigned ^{15}N -HSQC spectrum is shown in Fig. 2.

The coordinates of the 25 minimized conformers obtained with DYANA, on the basis of the NMR restraints, have been deposited in the Protein Data Bank under the accession number 1SS3. The resulting structures satisfied the experimental constraints with small deviations from the idealized covalent geometry, and most of the backbone torsion angles for non-glycine residues lay within the allowed regions in the Ramachandran plot. The statistics and details of quality and precision for the accepted structures are summarized in Table I, and a superposition of the selected structures is shown in Fig. 3. The root-mean-square deviation value for backbone heavy atoms of the most structured regions (excluding residues 1 and 2 from the N terminus and residues 34–50 from the unstructured C terminus) is 1.14 Å. The maximum violation is 0.29 Å, and the average sum of violations for the whole set of distance constraints is 4.0 Å.

The global fold of Ole e 6 consists of two nearly antiparallel α -helices, spanning residues 3–19 and 23–33, that are bound by a short loop and maintained together with the help of the three disulfide bonds. The average angle formed between these two helices is 140° in the 25 final minimized structures. Both α -helices have a very regular internal hydrogen bonding pattern. As expected, stretches of $\text{CO}(i)\text{--HN}(i+4)$ hydrogen bonds are

TABLE I
Summary of NMR structure calculations and structural statistics of the solution structure of rOle e 6

Constraints	
Distances	
Intraresidual	164
Sequential	153
Medium range ($i-j < 5$)	114
Long range ($i-j \geq 5$)	55
Total distances	486
Angles	
ϕ	24
Total	510
DYANA (50 structures)	
Target function	
Average	0.41
Minimal	0.34
Maximal	0.55
Maximal violation (Å)	
Average	0.21
Minimal	0.18
Maximal	0.29
Sum of violations (Å)	
Average	4.0
Minimal	3.5
Maximal	5.1
AMBER (25 structures)	
Mean energy (kcal/mol)	−284
Root-mean-square deviations from ideal geometry	
Bond lengths (Å)	0.00853
Bond angles (degrees)	1.990
Ramachandran plot	
Favorable regions	70.0%
Additional allowed	26.7%
Generously allowed	2.0%
Disallowed	1.3%
Averaged pairwise root-mean-square deviations (Å)	
Global	
Backbone	5.85 ± 2.64
Heavy atoms	6.86 ± 2.65
Structured region (3–33)	
Backbone	0.25 ± 0.10
Heavy atoms	1.14 ± 0.15
Helix I (3–19)	
Backbone	0.13 ± 0.05
Heavy atoms	0.98 ± 0.17
Helix II (23–33)	
Backbone	0.16 ± 0.06
Heavy atoms	1.20 ± 0.25

found in regions 3–20 and 22–33 in the resulting structures. The C-terminal end (residues 34–50) behaves as an unstructured tail. In addition to the mentioned hydrogen bonds participating in the secondary structure, we found an important one appearing in all conformers that involves the side chain of the histidine residue, His-13 H ϵ , and the carbonyl backbone atom of Phe-23. This hydrogen bond could be important for the stability of the protein as it keeps together residues in the two α -helices.

Many side chain conformations of residues in rOle e 6 are well defined, with 22 residues (5, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 21, 24, 25, 26, 27, 29, 30, 31, 32, and 33) having χ_1 circular variances lower than 0.05. These residues with well defined side chain conformations mainly correspond to residues exhibiting high values of surface inaccessibility and to residues located in the secondary structural elements. One important exception is Phe-23. This hydrophobic residue is at the N-terminal end of the second α -helix, and unexpectedly, its side chain is completely disordered and surface-exposed. During the NMR assignment process only a single chemical shift value could be determined for the aromatic protons, indicating the

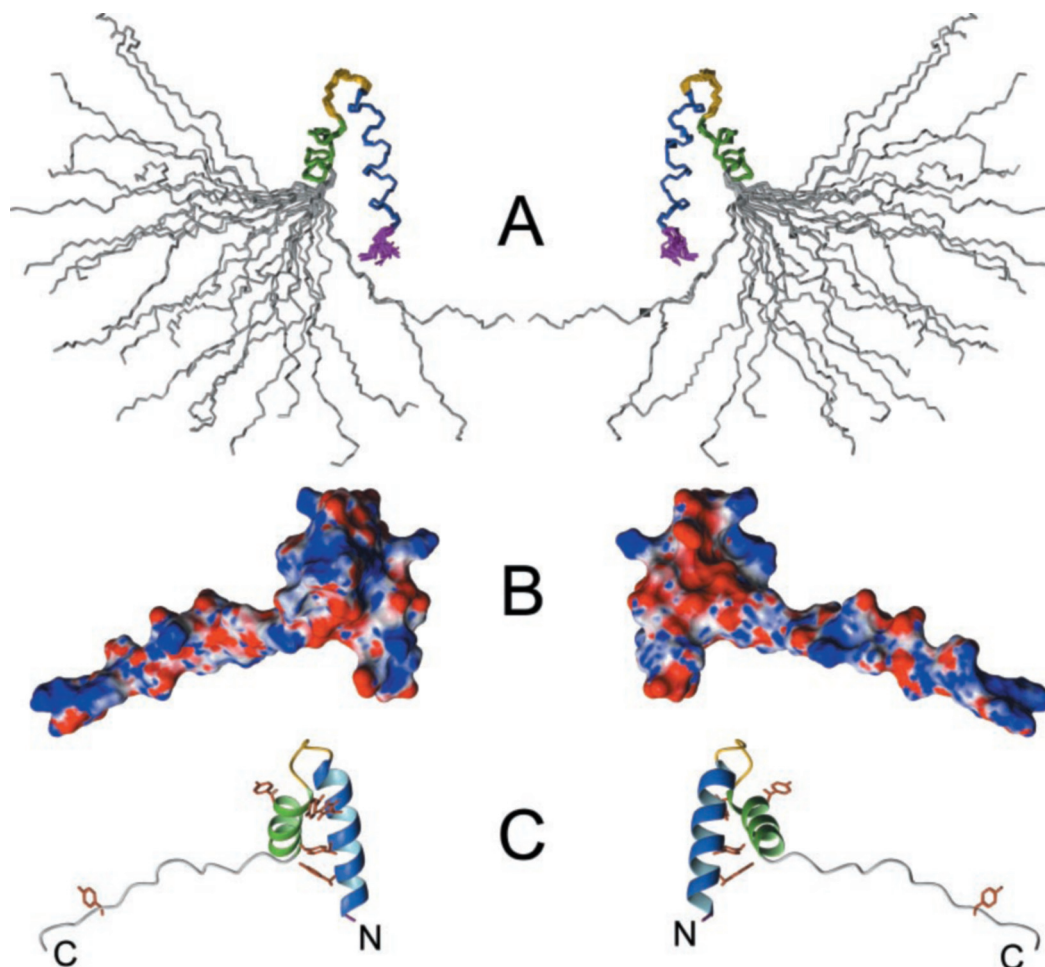


FIG. 3. **Details of the solution structure of Ole e 6.** A, superposition of 25 conformers. Helix I is colored in blue, helix II in green, the loop between helices in yellow, the N-terminal end in magenta, and the unstructured C terminus in gray. B, electrostatic surface. Positive areas are represented in blue, and negative areas are represented in red. C, ribbon diagram colored as in A. A, B, and C were produced with MOLMOL.

particular magnetic environment of this side chain. The degeneration of all ring proton chemical shifts in all the experimental conditions studied probably impeded the acquisition of a more precise set of distant restraints. In the future, additional data different from the NOE-based distance constraints, residual dipolar couplings, and others should be very useful to address whether this high disorder is compatible with the right conformation of Phe-23 in solution. Finally, those residues located on the long C-terminal tail are also highly disordered (see Table I and Fig. 3).

The hydrophobic core of this small protein has an unusual composition. Residues with accessible surface below 20% are Phe-5, Cys-8, Tyr-9, Cys-12, His-13, Cys-16, Ser-17, Cys-26, Cys-30, and Cys-34. Thus, with the exception of Ser-17, the core of the protein is composed only of aromatic residues and cysteines (Fig. 4). Their relative disposition is also very specific; all of the aromatic rings are in one face of the interhelix surface, whereas the three S–S bonds are just opposite in the other face. Additionally, each disulfide bond is located on the top of the inner face of one ring: disulfide 16–26 on the top of His-13, disulfide 12–30 on Tyr-9, and disulfide 8–34 on Phe-5 (Fig. 4).

Ole e 6 has a great number of exposed residues, most of them bearing charged groups (11 negative and 9 positive charges). The surface of the protein shows a well defined charge distribution. The unstructured tail is mainly positive, whereas the structured part has a positive and a negative face as can be seen in Fig. 3B. The main cluster of positive charges is com-

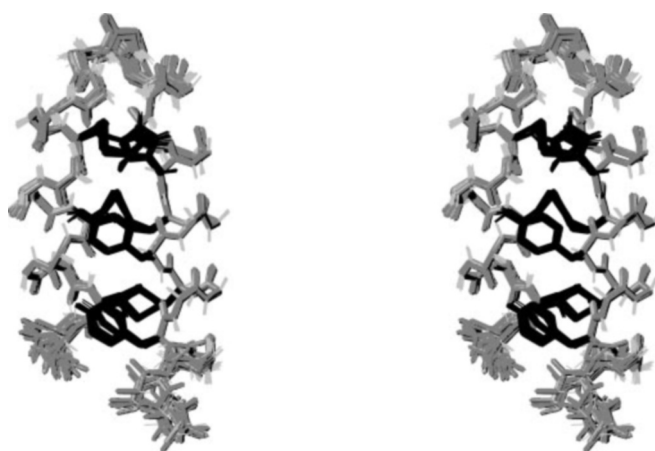


FIG. 4. **Stereoscopic representation of the protein core of Ole e 6.** The backbone atoms are in gray, and the aromatic groups and the disulfide bonds are in black.

posed of residues Lys-19 and Lys-29, and the negative clusters are composed of Glu-15, Asp-18, and Asp-33 in one side and Glu-27 and Asp-31 in the other side.

DISCUSSION

Ole e 6 is the first olive pollen allergen whose three-dimensional structure has been determined. No homologue protein was found for Ole e 6 in the data base banks, and only polypep-

TABLE II
Sequence similarity between more accessible fragments of Ole e 6 and other allergens

Underlined letters indicate identical residues, and boldface letters indicate substitution for similar residues. The peptides are ordered by their sequence similarity index as defined in the Structural Database of Allergenic Protein (SDAP) (56).

Matching region	Sequence description	Source	Reference
1 <u>DEAQFK</u> 6	Ole e 6	Olive pollen	25
76 <u>DHANFK</u> 81	Bet v 1	White birch pollen	62
18 <u>DKGNGFTF</u> 25	Ole e 6	Olive pollen	25
1 <u>DKGPGFVV</u> 8	Phl p 11	Ray grass pollen	63
26 <u>DQGQTWTV</u> 33	Cup a 3	Cypress pollen	Unpublished data
52 <u>DQGQTWTV</u> 59	Jun a 3	Mountain cedar pollen	64
35 <u>SVKDVKEKLENYKPKN</u> 50	Ole e 6	Olive pollen	25
63 <u>PFKYVKDRVDEVDHTN</u> 78	Bet v 1	White birch pollen	62
63 <u>PVKYVKERVEEIDHTN</u> 78	Car b 1	Hornbeam pollen	65
64 <u>PFKYVKERVEEVDHTN</u> 79	Cor a 1	Hazel leaf	66
244 <u>TNEDIAEKLESPDDER</u> 259	Glycinin	Soybean	67

tide sequences derived of two genomic structures were detected to have a degree of similarity. The alignment of the amino acid sequence of Ole e 6 with those deduced from genes of *N. tabacum* and *C. elegans* rendered an identity degree of 52 and 28%, respectively, and the Cys- X_3 -Cys- X_3 -Cys motifs were conserved. Unfortunately, no data about these putative proteins or about their function or structure are available so far.

The solution structure of the Ole e 6 allergen determined here consists of a well defined globular N-terminal body composed of two antiparallel α -helices and a long and unstructured C-terminal tail. The most characteristic feature of this structure is the internal core made up by the disulfide bridges 8–34, 12–30, and 16–26 and the aromatic rings of Phe-5, Tyr-9, and His-13. The presence of this histidine, a positively charged residue, in the hydrophobic core deserves some attention. The global accessible surface area of His-13 is very low (less than 9%), and specifically the accessibility of some ring atoms is less than 1%. It is well known that the presence of a charged residue in the protein interior is energetically unfavorable (43), and sometimes additional interactions are necessary to stabilize the local structure (44, 45). In Ole e 6, the NH ϵ of the imidazole ring participates in a hydrogen bond with the carbonyl group of Phe-23. This fact, together with the high burial, could contribute to decreasing the pK_a value of this group, making this histidine neutral at the natural physiological pH, diminishing the penalty of the electrostatic energy term, and increasing the stability of the protein.

On the other hand, the presence of a large amount of covalent, hydrophobic, and hydrogen bond interactions between the two helices suggests a high stability for this part of the protein. In fact, the absence of a two-state denaturation curve in the thermal transition from 20 to 85 °C and the extent of secondary elements that remained at high temperature indicate a great stability for the allergen, to which the disulphide bridges should also contribute. The presence of two grooves in the protein surface coincident with the two interhelix region faces is notable. Interestingly, the base of one of these grooves is composed of the outer faces of the aromatic rings of Phe-5, Tyr-9, His-13, and Phe-23. Their nearly planar disposition with respect to the surface suggests that they could interact with other groups as, for example, aromatic-aromatic or aromatic- π cation interactions. Additionally, the electrostatic surface shows at the N terminus well defined regions with positive and negative clusters. Most of the amino acids involved in such regions (Glu-15, Asp-18, Glu-27, Lys-29, Asp-31, and Asp-33) are conserved in the putative counterpart from tobacco, which suggests a possible biological role for these clusters in the activity of these pollen proteins. In summary, it seems that this protein has been designed to have a stable and structured core to specifically interact with other molecules, and its structure

points to the participation of electrostatic interactions as the driving forces for recognition.

Moreover, the unstructured tail can also play a biological and/or immunological role. The conformational chemical shifts of the residues in the C-terminal tail suggest that there is a tendency to form an α -helix in this region. Furthermore, estimation of secondary structure of Ole e 6 performed by theoretical methods and based on the amino acid sequence (46) gave an α -helix conformation for the segment at positions 34–47. It is likely that the C-terminal tail adopts a defined structure upon binding to a ligand or receptor or in different experimental conditions from those used here, as has been described for other proteins (47, 48).

At the present, vaccination is the most effective treatment to prevent type-I allergies. Ole e 6 seems to be a good candidate for the design of vaccines against olive pollen allergy as it is a relevant allergen with high prevalence among patients allergic to olive pollen. Thus, an exquisite definition of the epitopes is mandatory to develop these vaccines. In this sense, the correct definition of the epitopes in the molecule should be facilitated by the determination of the three-dimensional structure as the solvent accessibility of each region precisely indicates which areas are more susceptible to interact with the immunoglobulin (49, 50).

Crystallographic studies of Fab fragments of immunoglobulins complexed with their antigens have suggested that a contact surface covering approximately 600 Å² is necessary for the binding (51). Thus, from the results of this work, it can be deduced that the most exposed regions of Ole e 6 are (i) the highly flexible tail at the C-terminal end of the molecule (from residues 34 to 50, more than 2000 Å²), (ii) the loop connecting both helices and its vicinity (from Asp-18 to Phe-25, 715 Å²), and (iii) the first residues of the protein (from Asp-1 to Lys-6, 742 Å²). Other possible epitopes are those formed by residues located in the external faces of the helices with accessible surface areas of 789 Å² in helix I and 788 Å² in helix II. All of these regions could be important for choosing key residues in the recognition and, as a consequence, how to design hypoallergenic forms of the protein.

Structural comparison of proteins that induce an allergenic response can be used to predict allergenic cross-responses (52), eventually to determine possible characteristics of IgE recognition (50), and to predict the regions recognized by the immunoglobulins (10, 53, 54). Comparison of the whole sequence of Ole e 6 with the Swiss-Prot data base had shown only homology with a cysteine-rich putative protein from *N. tabacum* (55) of unknown three-dimensional structure. However, using the previously defined regions of Ole e 6 with high accessibility, a detailed search of homologues in the specific Structural Database of Allergenic Protein (SDAP) (56) has shown similarities with fragments of some allergens (Table II). Among them, the

only one with a known three-dimensional structure is Bet v 1 (1BTv) (57). The peptide homologue of the C-terminal tail of Ole e 6 in Bet v 1 is part of a β -sheet, and it is partially accessible. The NMR data based on the conformational chemical shifts and secondary structure prediction suggest a low natural tendency of the fragment 35–50 of Ole e 6 to adopt an α -helix conformation. However, we cannot discard the possibility that it could adopt a conformation similar to that of the sequence 63–78 from Bet v 1 in the presence of a ligand and/or the immunoglobulin. Like the C-terminal region, the N-terminal region (residues 1–6) of Ole e 6 shows homology with a peptide from Bet v 1 located on a loop (residues 76–81). Finally, the comparison between the sequence of the loop of Ole e 6 (from position 18 to position 25) and other pollen allergens revealed a low but significant identity with segments of Phl p 11, Cup a 3, and Jun a 3, but no structures have been determined to allow a comparison of shapes or accessibilities. However, all of these detected similarities seem likely to be too low to be responsible for a significant cross-reactivity between Ole e 6 and these allergens. In any case, Ig binding inhibition experiments with purified allergens should be performed to clarify the possible cross-reactivity between them.

The biological function of Ole e 6 is still unknown. To try to shed some light on this point, we have compared Ole e 6 with other published protein structures using the DALI server (58). The comparison reveals homology with protein fragments holding two α -helices with similar relative orientation to those in Ole e 6. It is interesting to note that similarities have been found with protein fragments that are part of channels, such as ATP synthase (59) and the mechanosensitive channel MscS (60) from *E. coli*, and with a protein that interacts with lipids, saposin B (61) from humans. As mentioned above, the specific location of aromatic residues in the α -helix surfaces of Ole e 6, and the existence of two well defined basic and acid clusters, and their conservation on the putative counterpart from tobacco pollen, allow us to suggest a possible involvement of these structural motifs in the biochemical activity of the protein.

In conclusion, we have obtained the first well defined structure of an allergen from olive pollen, one of the most important causes of pollinosis in Mediterranean countries. The information derived from this structure can be used to simplify the mapping of the epitopes of the protein in order to design new diagnostic and therapeutic tools and to investigate the cellular activity of this pollen protein.

REFERENCES

- Venarske, D., and deShazo, R. D. (2003) *South. Med. J.* **96**, 1049–1054
- Miescher, S. M., and Vogel, M. (2002) *Mol. Aspects Med.* **23**, 413–462
- Budde, I. K., and Aalberse, R. C. (2003) *Clin. Exp. Allergy* **33**, 1175–1182
- Gould, H. J., Sutton, B. J., Beavil, A. J., Beavil, R. L., McCloskey, N., Coker, H. A., Fear, D., and Smurthwaite, L. (2003) *Annu. Rev. Immunol.* **21**, 579–628
- Valenta, R., Lidholm, J., Niederberger, V., Hayek, B., Kraft, D., and Gronlund, H. (1999) *Clin. Exp. Allergy* **29**, 896–904
- Suphioglu, C., Blaher, B., Rolland, J. M., McCluskey, J., Schappi, G., Kenrick, J., Singh, M. B., and Knox, R. B. (1998) *Mol. Immunol.* **35**, 293–305
- Jarvinen, K. M., Chatchatee, P., Bardina, L., Beyer, K., and Sampson, H. A. (2001) *Int. Arch. Allergy Immunol.* **126**, 111–118
- Mine, Y., and Wei Zhang, J. (2002) *Biochem. Biophys. Res. Commun.* **292**, 1070–1074
- Bohle, B., Radakovics, A., Jahn-Schmid, B., Hoffmann-Sommergruber, K., Fischer, G. F., and Ebner, C. (2003) *Eur. J. Immunol.* **33**, 3303–3310
- Furmonaviciene, R., Sewell, H. F., and Shakib, F. (2000) *Clin. Exp. Allergy* **30**, 1307–1313
- D'Amato, G., Spiekma, F. T., Liccardi, G., Jager, S., Russo, M., Kontou-Fili, K., Nikkels, H., Wuthrich, B., and Bonini, S. (1998) *Allergy* **53**, 567–578
- Bousquet, J., Cour, P., Guerin, B., and Michel, F. B. (1984) *Clin. Allergy* **14**, 249–258
- Bousquet, J., Guerin, B., Hewitt, B., Lim, S., and Michel, F. B. (1985) *Ann. Allergy* **69**, 439–440
- Kernerman, S. M., McCullough, J., Green, J., and Ownby, D. R. (1992) *Ann. Allergy* **69**, 493–496
- Martín-Orozco, E., Cárdbaba, B., del Pozo, V., de Andrés, B., Villalba, M., Gallardo, S., Rodríguez-García, M. I., Fernández, M. C., Alché, J. D., Rodríguez, R., Palomino, P., and Lahoz, C. (1994) *Int. Arch. Allergy Immunol.* **104**, 160–170
- Pajaron, M. J., Vila, L., Prieto, I., Resano, A., Sanz, M. L., and Oehling, A. K. (1997) *Allergy* **52**, 829–835
- Wahl, R., Schmid-Grendelmeier, P., Cromwell, O., and Wuthrich, B. (1996) *J. Allergy Clin. Immunol.* **98**, 99–106
- González, E., Villalba, M., and Rodríguez, R. (2001) *Clin. Exp. Allergy* **31**, 313–321
- Obispo, T. M., Melero, J. A., Carpizo, J. A., Carreira, J., and Lombardero, M. (1993) *Clin. Exp. Allergy* **23**, 311–316
- Rodríguez, R., Villalba, M., Batanero, E., González, E. M., Monsalve, R. I., Huecas, S., Tejera, M. L., and Ledesma, A. (2002) *Allergy* **57**, 6–16
- Rodríguez, R., Villalba, M., Monsalve, R. I., and Batanero, E. (2001) *Int. Arch. Allergy Immunol.* **125**, 185–195
- Barral, P., Tejera, M. L., Treviño, M. A., Villalba, M., Bruix, M., and Rodríguez, R. (2004) *Protein Expression Purif.*, in press
- Huecas, S., Villalba, M., and Rodríguez, R. (2001) *J. Biol. Chem.* **276**, 27959–27966
- Barral, P., Batanero, E., Palomares, O., Quiralte, J., Villalba, M., and Rodríguez, R. (2004) *J. Immunol.* **172**, 3644–3651
- Batanero, E., Ledesma, A., Villalba, M., and Rodríguez, R. (1997) *FEBS Lett.* **410**, 293–296
- Huecas, S., Villalba, M., González, E., Martínez-Ruiz, A., and Rodríguez, R. (1999) *Eur. J. Biochem.* **261**, 539–546
- Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) *J. Biomol. NMR* **6**, 135–140
- Aue, W. P., Bartholdi, E., and Ernst, R. R. (1976) *J. Chem. Phys.* **64**, 2229–2246
- Bax, A., and Davis, D. G. (1985) *J. Magn. Reson.* **65**, 355–360
- Kumar, A., Ernst, R. R., and Wuthrich, K. (1980) *Biochem. Biophys. Res. Commun.* **95**, 1–6
- Bodenhausen, G., and Ruben, D. J. (1980) *Chem. Phys. Lett.* **69**, 185–189
- Palmer, A. G., Cavanagh, J., Byrd, R. A., and Rance, M. (1992) *J. Magn. Reson.* **96**, 416–424
- Crowley, P., Ubbink, M., and Otting, G. (2000) *J. Am. Chem. Soc.* **122**, 2968–2969
- Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6**, 277–293
- Helgstrand, M., Kraulis, P., Allard, P., and Hard, T. (2000) *J. Biomol. NMR* **18**, 329–336
- Güntert, P., Braun, W., and Wuthrich, K. (1991) *J. Mol. Biol.* **217**, 517–530
- Güntert, P., Mumenthaler, C., and Wuthrich, K. (1997) *J. Mol. Biol.* **273**, 283–298
- Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., III, Wang, J., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Tsui, V., Gohlke, H., Radmer, R. J., Duan, Y., Pitera, J., Massova, I., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (2002) AMBER 7, University of California, San Francisco
- Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graph.* **14**, 51–55
- Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8**, 477–486
- Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., New York
- Seavey, B. R., Farr, E. A., Westler, W. M., and Markley, J. L. (1991) *J. Biomol. NMR* **1**, 217–236
- Dao-pin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W., and Matthews, B. W. (1991) *Biochemistry* **30**, 11521–11529
- Burley, S. K., and Petsko, G. A. (1985) *Science* **229**, 23–28
- Ma, J. C., and Dougherty, D. A. (1997) *Chem. Rev.* **97**, 1303–1324
- Garnier, J., Gibrat, J. F., and Robson, B. (1996) *Methods Enzymol.* **266**, 540–553
- Reyda, S., Sohn, C., Klebe, G., Rall, K., Ullmann, D., Jakubke, H. D., and Stubbs, M. T. (2003) *J. Mol. Biol.* **325**, 963–977
- Dirr, H. W., and Wallace, L. A. (1999) *Biochemistry* **38**, 15631–15640
- Breiteneder, H., and Ebner, C. (2001) *Curr. Opin. Allergy Clin. Immunol.* **1**, 261–267
- Soman, K. V., Midoro-Horiuti, T., Ferreón, J. C., Goldblum, R. M., Brooks, E. G., Kurosky, A., Braun, W., and Schein, C. H. (2000) *Biophys. J.* **79**, 1601–1609
- Davies, D. R., Padlan, E. A., and Sheriff, S. (1990) *Annu. Rev. Biochem.* **59**, 439–473
- Ipsen, H., and Lowenstein, H. (1997) *Clin. Rev. Allergy Immunol.* **15**, 389–396
- Lascombe, M. B., Gregoire, C., Poncet, P., Tavares, G. A., Rosinski-Chupin, I., Rabillon, J., Goubran-Botros, H., Mazie, J. C., David, B., and Alzari, P. M. (2000) *J. Biol. Chem.* **275**, 21572–21577
- Roy, D., Ghosh, D., and Gupta-Bhattacharya, S. (2003) *Biochem. Biophys. Res. Commun.* **307**, 422–429
- Cheung, A. Y., Wu, H. M., Di Stilio, V., Glaven, R., Chen, C., Wong, E., Ogdhal, J., and Estavillo, A. (2000) *Ann. Bot. (Lond.)* **85**, 29–37
- Ivancic, O., Schein, C. H., and Braun, W. (2002) *Bioinformatics (Oxf.)* **18**, 1358–1364
- Gajhede, M., Osmark, P., Poulsen, F. M., Ipsen, H., Larsen, J. N., Joost van Neerven, R. J., Schou, C., Lowenstein, H., and Spangfort, M. D. (1996) *Nat. Struct. Biol.* **3**, 1040–1045
- Holm, L., and Sander, C. (1995) *Trends Biochem. Sci.* **20**, 478–480
- Rastogi, V. K., and Girvin, M. E. (1999) *Nature* **402**, 263–268
- Bass, R. B., Strop, P., Barclay, M., and Rees, D. C. (2002) *Science* **298**, 1582–1587
- Ahn, V. E., Faull, K. F., Whitelegge, J. P., Fluharty, A. L., and Prive, G. G. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 38–43

62. Swoboda, I., Jilek, A., Ferreira, F., Engel, E., Hoffmann-Sommergruber, K., Scheiner, O., Kraft, D., Breiteneder, H., Pittenauer, E., Schmid, E., Vicente, O., Heberle-Bors, E., Ahorn, H., and Breitenbach, M. (1995) *J. Biol. Chem.* **270**, 2607–2613
63. Marknell DeWitt, A., Niederberger, V., Lehtonen, P., Spitzauer, S., Sperr, W. R., Valent, P., Valenta, R., and Lidholm, J. (2002) *Clin. Exp. Allergy* **32**, 1329–1340
64. Midoro-Horiuti, T., Goldblum, R. M., Kurosky, A., Wood, T. G., and Brooks, E. G. (2000) *J. Immunol.* **164**, 2188–2192
65. Larsen, J. N., Stroman, P., and Ipsen, H. (1992) *Mol. Immunol.* **29**, 703–711
66. Luttkopf, D., Muller, U., Skov, P. S., Ballmer-Weber, B. K., Wuthrich, B., Skamstrup Hansen, K., Poulsen, L. K., Kastner, M., Haustein, D., and Vieths, S. (2002) *Mol. Immunol.* **38**, 515–525
67. Momma, T., Negoro, T., Hirano, H., Matsumoto, A., Udaka, K., and Fukazawa, C. (1985) *Eur. J. Biochem.* **149**, 491–496