Pan, L. P., Durham, B., Wolinscka, J., & Millett, F. S. (1988) *Biochemistry 27*, 7180-7184.
Smith, H. T., Staudenmeyer, N., & Millett, F. (1977) *Biochemistry 16*, 4971-4975.
Sprintschnik, G. H. W., Dirsch, P. P., & Whitten, D. G. (1977) *J. Am. Chem. Soc. 99*, 4947.
Sutin, N. (1982) *Acc. Chem. Res. 15*, 275-282.
Sutin, N., & Creutz, C. (1978) *Adv. Chem. Ser. 168*, 1-27.
Swanson, T., Trus, B. L., Mandel, N., Mandel, G., Kallai, O., & Dickerson, R. E. (1977) *J. Biol. Chem. 252*, 759-770.
Winkler, J. R., Nocera, D. G., Yocom, K. B., Bordignon, E., & Gray, H. B. (1982) *J. Am. Chem. Soc. 104*, 5798-5800.

Complete Primary Structure of a *Lolium perenne* (Perennial Rye Grass) Pollen Allergen, *Lol p* III: Comparison with Known *Lol p* I and II Sequences†

Aftab A. Ansari,*,‡ Ponniah Shenbagamurthi,§ and David G. Marsh‡

Clinical Immunology Division, Department of Medicine, The Johns Hopkins University School of Medicine, 5601 Loch Raven Boulevard, Baltimore, Maryland 21239, and Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received May 2, 1989; Revised Manuscript Received June 27, 1989

ABSTRACT: The complete amino acid sequence of a *Lolium perenne* (rye grass) pollen allergen, *Lol p* III, determined by the automated Edman degradation of the protein and its selected fragments, is reported in this paper. Cleavage by enzymatic and chemical techniques established unambigusously the sequence for this 97-residue protein (M_r = 10909), which lacks cysteine and shows no evidence of glycosylation. The sequence of *Lol p* III is very similar to that of another *L. perenne* allergen, *Lol p* II, which was sequenced recently; of the 97 positions in the two proteins, 57 are occupied by identical amino acids (59% identity). In addition, both allergens share a similar structure with an antibody-binding fragment of a third *L. perenne* allergen, *Lol p* I. Since human antibody responsiveness to all these three allergens is associated with HLA-DR3, and since the structure common to the three molecules shows high degrees of amphipathicity in *Lol p* II and III, we speculate that this common segment in the three molecules might contain or contribute to the respectively Ia/T-cell sites.

Protein antigens contain two classes of epitopes: B-cell epitopes, which are recognized by specific antibodies, and Ia/T-cell epitopes, which are recognized by an MHC[1] (Ia) molecule and a T-cell receptor. In vivo, the antigen molecule is cleaved into smaller fragments inside an antigen presenting cell (APC), such as a macrophage, a dendritic cell, a B cell, etc.; one or more fragments containing Ia/T-cell epitope(s) bind to an Ia molecule, and this antigen-Ia complex is presented by the APC to a specific T-cell receptor on a helper T cell (Schwartz, 1985). The T cell is stimulated to proliferate clonally and secrete interleukins, which activate B cells to proliferate clonally, differentiate, and produce antibodies specific for the B-cell epitopes of the antigen. An understanding of the genetic control of the human immune response requires delineating the structures (Ia/T-cell sites) in an antigen molecule that are presented to the T cells by Ia molecules. We have been studying the genetics of human immune responsiveness using the *Lolium perenne* (rye grass) pollen allergens *Lol p* I, II, and III as model antigens.

Lol p I (Rye I), the major allergen of *L. perenne* pollen, is a glycoprotein (M_r = 27000-32000) toward which 85-90% of grass-allergic patients are sensitive (Freidhoff et al., 1986); *Lol p* II and III are proteins (M_r = 11000 in both cases) toward which 45% of grass-allergic patients are sensitive in each case (Freidhoff et al., 1986; Ansari et al., unpublished observation). Immunologically, *Lol p* II and III are cross-reactive with each other, but neither shows any cross-reactivity with *Lol p* I (Ansari et al., 1987). All of these allergens exist in multiple "isoallergenic" forms which are variants but appear to be immunologically indistinguishable (Johnson & Marsh, 1965a,b). *Lol p* III is a group of basic proteins with isoelectric points in the range of 9.0-9.4, with the main isoallergenic forms *Lol p* IIIB (p*I* = 9.0) and *Lol p* IIIA (p*I* = 9.4), the former being the more abundant quantitatively.

The *Lol p* allergens have proved important in studies concerning the genetics of human immune responsiveness (Freidhoff et al., 1988; Ansari et al., 1989a,b). We have found that human IgE and IgG Ab responsiveness to *Lol p* I-III is significantly associated with HLA-DR3; in the case of *Lol p* III the association extends to DR5 also. The partial sequence

† This work was supported by Grants AI-19727 and AI-25372 from the National Institutes of Health and by Johns Hopkins Institutional Grant M-53-1007. This is Publication No. 764 from the O'Neill Laboratories, The Good Samaritan Hospital, Baltimore, MD.

* Address correspondence to this author at the Clinical Immunology Division, Department of Medicine, The Johns Hopkins University School of Medicine, 5601 Loch Raven Blvd., Baltimore, MD 21239.
‡ Department of Medicine, The Johns Hopkins University School of Medicine.
§ Department of Biological Chemistry, The Johns Hopkins University School of Medicine.

[1] Abbreviations: Ab, antibody; HLA, human leucocyte antigen; HPLC, high-performance liquid chromatography; rpHPLC, reverse-phase HPLC; *Lol p* I, II, and III, designations given under the IUIS Allergen Nomenclature System (Marsh et al., 1986) for *Lolium perenne* pollen allergens previously known as Rye I, II, and III, respectively; MHC, major histocompatibility complex; RIA, radioimmunoassay; TFA, trifluoroacetic acid; TPCK, tosylphenylalanine chloromethyl ketone; TLCK, tosyllysine chloromethyl ketone.

0006-2960/89/0428-8665$01.50/0

of *Lol p* I is known (Cottam et al., 1986; Esch & Klapper, 1989), and the complete sequence of *Lol p* II was recently reported from this laboratory (Ansari et al., 1989c). In this paper, we have determined the complete primary structure of *Lol p* III and have compared this sequence with the *Lol p* I and II sequences. We performed theoretical analyses to predict B-cell sites and Ia/T-cell sites in the protein. Finally, we attempted to correlate these findings with what we know of the human immune response to these allergens.

Materials and Methods

Lol p III. The major, less basic form of *Lol p* III, *Lol p* IIIB (p*I* = 9.0), purified as described (Ansari et al., 1987), was the same preparation as used in our previous studies (Ansari et al., 1987, 1989a,b). On agarose electrophoresis, the *Lol p* IIIB preparation gave a single well-defined band. On SDS–PAGE, the protein gave one bond corresponding to its reported molecular weight of 11 000. Isoelectricfocusing of the *Lol p* III preparation gave a diffuse band which did not focus properly in repeated experiments. No contamination by *Lol p* I or II was detectable by Coomassie Blue staining following agarose electrophoresis, SDS–PAGE, or isoelectric focusing.

Proteolytic Enzymes. The following enzymes were from Sigma Chemical Co.: TPCK-treated trypsin from bovine pancreas (T8642, lot no. 55F-8065), TLCK-treated α-chymotrypsin from bovine pancreas (C3142, lot no. 75F-8025), staphylococcal V8 protease (P8400, lot no. 16F-03271), and clostripain (C7403, lot no. 37F-68422). Endoproteinase ASP-N was a kind gift from Dr. E. Appella of the NIH, Bethesda, MD.

Citraconylation and Decitraconylation. Lol p III was citraconylated as described earlier (Ansari et al., 1976) in order to limit the tryptic digestion to arginine peptide bonds. After the digestion, the citraconyl groups were removed by suspending the digest in 0.1 M pyridine–acetate buffer, pH 3.5, for 24 h at 25 °C.

Tryptic Digestion. Lol p III (16 nmol) was citraconylated before digestion in 10 mM ammonium bicarbonate, pH 7.8, with TPCK-treated trypsin from bovine pancreas (Sigma T8642, batch no. 55F-8065). The enzyme-to-substrate weight ratio was 1:100, and the digestion was carried out at 37 °C for 90 min. After the digestion, the citraconyl groups were removed by suspending the digest in 0.1 M pyridine–acetate buffer, pH 3.5, for 24 h at 25 °C. The mixture was lyophilized and then reconstituted in water containing 0.05% TFA. Some peptides were lost as precipitates.

Chymotryptic Digestion. Lol p III (9 nmol) was digested in 10 mM ammonium bicarbonate, pH 7.8, with TLCK-treated α-chymotrypsin from bovine pancreas (Sigma C3142, batch no. 75F-8025). The enzyme-to-substrate weight ratio was 1:50, and the digestion was carried out at 37 °C for 18 h. The mixture was lyophilized and then reconstituted in water containing 0.05% TFA. No insoluble material was noticed.

Digestion with Glu-Specific Staphylococcal V8 Protease. Lol p III (12 nmol) was digested in 0.1 M sodium phosphate/2 mM EDTA, pH 7.8, with the V8 protease (Sigma) at 37 °C for 18 h by using an enzyme-to-substrate ratio of 1:30. The digest contained some insoluble material. No lyophilization was performed before HPLC.

Digestion with Endoproteinase ASP-N. Lol p III (12 nmol) was digested in 0.05 M ammonium bicarbonate, pH 7.8, with endoproteinase ASP-N. The enzyme-to-substrate ratio was 1:100 (w/w), and the reaction was carried out at 25 °C for 18 h. In one experiment, *Lol p* III (2.6 nmol) was digested with the enzyme in 2 M deionized urea at an enzyme-to-substrate ratio of 1:80 (w/w) for 22 h at 37 °C.

Clostripain Digestion of Lol p III. Clostripain (Sigma C7403, batch no. 37F-68422) was activated for 2–3 h at 25 °C by dissolving the enzyme in 2.5 mM DTT/1.0 mM calcium chloride. *Lol p* III (16 nmol) was dissolved in 150 μL of 0.1 M sodium phosphate/1 mM DTT, pH 7.7. Activated clostripain was added (1:100 weight ratio) and the digestion carried out at 25 °C for 3 h. The mixture was diluted with 0.1% TFA in water and directly used in HPLC.

CNBr Digestion. To a solution of *Lol p* III (4.5 nmol) in 100 μL of 50 mM ammonium bicarbonate, pH 7.8, was added 5 μL of 100 mM DTT. The tube was flushed with nitrogen and incubated in the dark for 30 min. CNBr (1 mg) solution in 105 μL of 0.2 M HCl was added, the tube was flushed with nitrogen again, and the reaction was allowed to proceed at 25 °C for 14 h. The solution was diluted with 200 μL of 0.1% TFA in water and applied directly on the HPLC column.

Peptide Purification. Peptides were purified by rpHPLC using Waters C18 columns. Experimental details are presented with individual chromatograms. HPLC effluents were monitored at 214 nm; in a later part of the study, the effluents were also monitored at 280 nm in order to discriminate between peptides containing different proportions of aromatic amino acids.

Nomenclature of the Peptides. Peptides are designated by a serial number prefixed by a letter. The letters indicate the type of digestion: T, trypsin; C, chymotrypsin; E, Glu-specific protease; D, Asp-specific protease; CL, clostripain; CN, CNBr. The numbers in the peptide designations do not correspond to the order of their elution in HPLC, but rather to their position in the protein sequence, starting from the amino terminus. This system of nomenclature is followed for all those peptides for which data are presented. There are some peptides that were sequenced but for which data are not presented in detail, as these peptides cover regions of the proteins that are already sequenced by using the numbered peptides. The peptides for which sequence analysis data are not presented are indicated in the HPLC profiles (Figure S-2 of the supplementary material), in which the corresponding peaks carry a designation showing the amino acid residue positions spanned by the peptides.

Amino Acid Analysis. This was performed by using the Pico-Tag amino acid analysis system of Waters–Millipore according to the manufacturer's directions.

Amino Acid Sequence Analysis. Automated Edman degradation of peptides was performed on a Model 470A gas-phase protein sequencer (Applied Biosystems) using the 03RPTH program. BioBrene Plus (3 mg) was applied onto the trifluoroacetic acid etched glass fiber filter and subjected to three precycles of Edman degradation prior to sample application. The thiazolinone derivatives were converted to phenylthiohydantoin amino acids by using 25% TFA. The PTH-amino acid derivatives were separated, on-line, by rpHPLC on a PTH C18 column (2.1 × 220 mm; 5 μm; Applied Biosystems) using a gradient of acetonitrile and sodium acetate/trimethylamine buffer, containing 5% tetrahydrofuran on an Applied Biosystems Model 120A PTH analyzer.

Results

NH_2-Terminal Sequence of Lol p III. The *Lol p* III preparation (130 pmol) was initially subjected to the gas-phase sequence analysis for 25 cycles. A single, homogeneous sequence was obtained throughout, except for cycle 6, which showed two alternate amino acids, threonine and methionine, in approximately 4:1 ratio (the yield data can be found in Figure S-1 of the supplementary material). The presence of

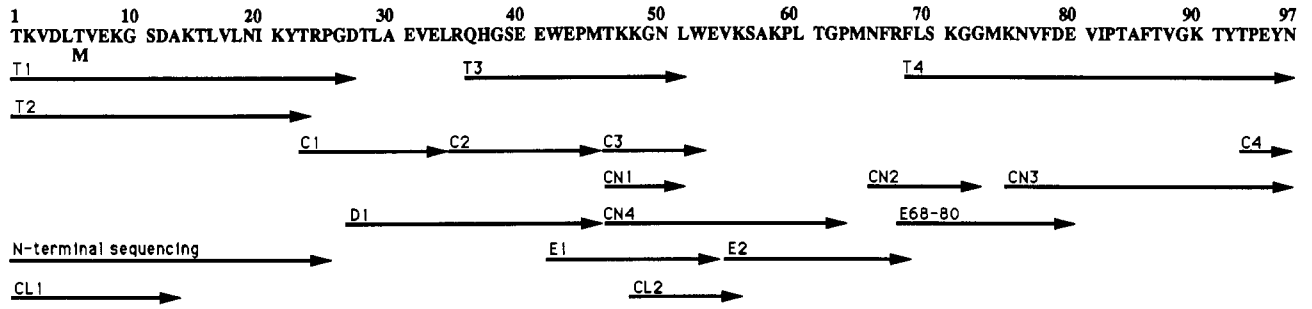

FIGURE 1: Amino acid sequence of *Lol p* IIIB and schematic outline of the data supporting the sequence. Amino acid residues are given in the single-letter code. Peptides obtained by cleavage with trypsin (T), chymotrypsin (C), proteinase ASP-N (D), *Staphylococcus aureus* V8 protease (E), clostripain (CL), and CNBr (CN) are indicated, with numbering from the NH_2 terminus (T1, T2, etc.). Residues identified by Edman degradation and the direction of sequence analysis are indicated by arrows. One polymorphic position (M substituting for T) was noted at position 6 as indicated. The amino acid composition, in mol/mol, calculated from the above main sequence (and from amino acid analysis of the *Lol p* III preparation, in parentheses) is as follows: D + N = 4 + 5 (9.69); E + Q = 9 + 1 (10.83); S = 4 (4.16); G = 8 (8.87); H = 1 (0.86); R = 3 (3.23); T = 11 (10.31); A = 4 (4.44); P = 6 (6.19); Y = 3 (3.00); V = 8 (7.63); M = 3 (3.52); C = 0 (0.00); I = 1 (1.29); L = 8 (8.21); F = 4 (3.21); K = 11 (9.54); W = 2 (not done).

these alternate residues in *Lol p* III variants was confirmed by isolating and sequencing peptides that showed one or the other of the two amino acids at this position, as described latter in this section.

Peptide Fragments of Lol p III. Peptide fragments were obtained by HPLC of various digests of *Lol p* III (the HPLC profiles can be found in Figure S-2 of the supplementary material). Most of the major peaks, and some minor peaks (a total of 41 peaks), were subjected to the automated gas-phase sequence analysis procedure (the yield data on 17 selected peptides are shown in Figure S-1 of the supplementary material). Additional peptides from each digestion were sequenced, but data on them are not shown (except in the chromatographic profiles, Figure S-2 of the supplementary material), as these peptides turned out to cover regions that were already sequenced by using one or more peptides. These additional peptides gave sequences identical with those included within the sequences presented. Because of this replication in sequence analysis, each region of the protein was ultimately confirmed by more than one, and sometimes up to four, independent sequence analyses.

Reconstruction of the Amino Acid Sequence. Figure 1 depicts the complete amino acid sequence of *Lol p* III, indicating the alignment of the peptides. This sequence is based on extensive overlaps in all the regions, each overlap being confirmed by sequence data from at least two peptides (data for all the replicates are not presented for the sake of brevity). Amino acid compositional data are consistent with the sequence (see legend to Figure 1).

Lol p III derived tryptic peptides T1–T4 yielded sequence information on 74% of the protein. The gap between T1 and T3 was filled by two chymotryptic peptides C1 and C2, and the alignment in this region was confirmed by a peptide D1 obtained by endoproteinase ASP-N digestion of the protein. Three peptides, namely, CL2 (from clostripain digestion), E2 (from V8 digestion), and CN2 (from CNBr digestion), filled the gap between T3 and T4 and confirmed the alignment in this region. Further confirmation of the alignment between E1 and E2 was provided by the peptide CL2. The carboxy-terminal sequence of *Lol p* III was confirmed by the chymotryptic peptide C4, which gave the following sequence: TPEYN. Acid hydrolysis and amino acid analysis of this peptide yielded only these five amino acids, and the proportion of these amino acids with respect to each other was close to unity (Table I).

Difficult Regions. In *Lol p* III, the major problem was finding a peptide that would give an adequate overlap in the region around position 53 (Figure 1). The tryptic peptide T3, which was isolated in a low yield, could not be sequenced beyond Leu_{51}. The chymotryptic peptide C3 stopped at Trp_{52}. The peptide E1 stopped at Glu_{53}, while another peptide E2 started at Val_{54}. Clostripain digestion was carried out with a view to obtaining a peptide 36–66 resulting from cleavages at arginine residues. On the basis of absorbance at 280 nm and 214 nm, three peaks were selected for sequence analysis. One of them, CL1, was from the NH_2-terminus; therefore, the sequence analysis was stopped after 13 cycles. Another peptide (data not shown) was the result of cleavage at Arg_{35} and an unexpected cleavage at Lys_{48}. The next peptide CL2, a product of two unexpected cleavages at Lys_{47} and Lys_{55}, provided the desired overlap to E1 and E2. However, this was only a two-residue overlap. A CNBr digestion of *Lol p* III was carried out in the presence of DTT in the dark with a nitrogen flush, with a view to obtaining the fragment between the two methionine residues at 45 and 64. In an earlier CNBr cleavage of the protein in the absence of the reducing agent, such a peptide could not be obtained as a result of cleavage of the tryptophanyl bond at position 52. In the second attempt (with the reducing agent), the desired peptide CN4 was obtained, which further confirmed the overlap between E1 and E2.

Table I: Confirmation of the Carboxy-Terminal Sequence of *Lol p* III Using Peptide C4

amino acid	AA analysis pmol	AA analysis mol/mol	sequence analysis (mol/mol)
Thr	140	1.06	1
Pro	146	1.11	1
Glx	135	1.02	1 (Glu)
Tyr	139	1.05	1
Asx	100	0.76	1 (Asn)

Variant of Lol p III. *Lol p* III was found to show an alternate amino acid at position 6. NH_2-terminal sequence analysis of the undigested protein showed Thr and Met at cycle 6 in the ratio 47/11 (the yield data can be seen in Figure S-1 of the supplementary material). Most *Lol p* III peptides, including T1 and T2 (and five additional peptides for which data are not shown), showed Thr at this position, but clostripain peptide CL1 showed Met at cycle 6.

Specificity of Enzymic and Chemical Cleavages. It is interesting to make a note of the unexpected peptide bond cleavages observed during the course of this investigation. Table II lists those cleavages that are usually not expected from the specificities of the respective reagents. The table shows

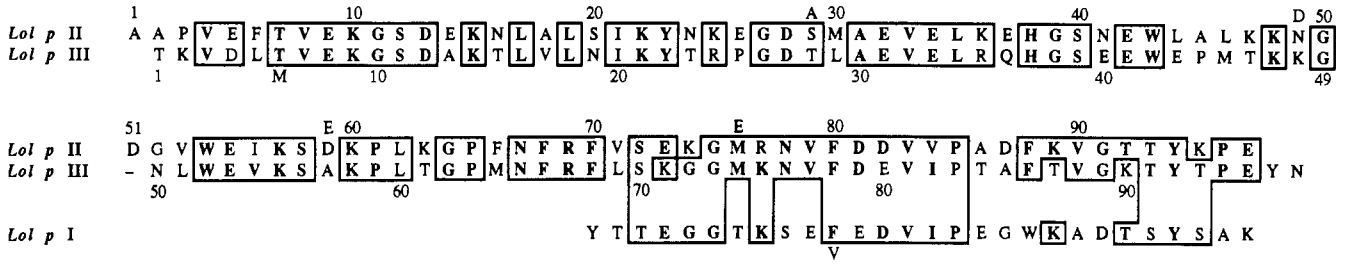

FIGURE 2: Alignment of amino acid sequence of *Lol p* IIIB with the complete sequence of *Lol p* II and an antibody-binding fragment of *Lol p* I. The *Lol p* III sequence is from this study, the *Lol p* II sequence is from Ansari et al. (1989c), and the *Lol p* I sequence is from Esch and Kapper (1989). Variants found at certain positions for *Lol p* II are shown above its sequence, and those for *Lol p* III and I are shown below their sequences. Identical residues are shown in bold letters. These residues and other similar residues in the two sequences are indicated by boxes (the following residues were considered similar: E = D, K = R, S = T, and V = I).

Table II: Unexpected Peptide Bond Cleavages Observed

enzyme/chemical	unexpected bond cleaved	position in protein	yield[a] (%)
chymotrypsin	K-G[b]	48–49	5
	K-S[b]	55–56	69
	E-V	53–54	23
clostripain	K-K	47–48	>90
staphylococcal V8 protease	L-V	16–17	38
	W-E	42–43	0.4
	T-K	46–47	13
	R-F	67–68	9
endoproteinase ASP-N	L-N	18–19	4
(*Pseudomonas fragi*)	K-Y	21–22	8
CNBr	T-A	84–85	0.02
	T-Y	91–92	0.02

[a] Yield is expressed relative to the yield (based on sequence analysis) of the most abundant peptide in each hydrolysis. [b] These cleavages could be the result of trypsin contamination in the "TLCK-treated" chymotrypsin preparation used (Sigma C3142, lot no. 75F-8025).

the bond cleaved, position in the protein, and yield of the peptide relative to the yield (based on sequence analysis data) of the most abundant peptide in the group; this latter information gives an idea of whether the "unexpected" cleavage occurred at a low or high yield relative to an "expected" cleavage.

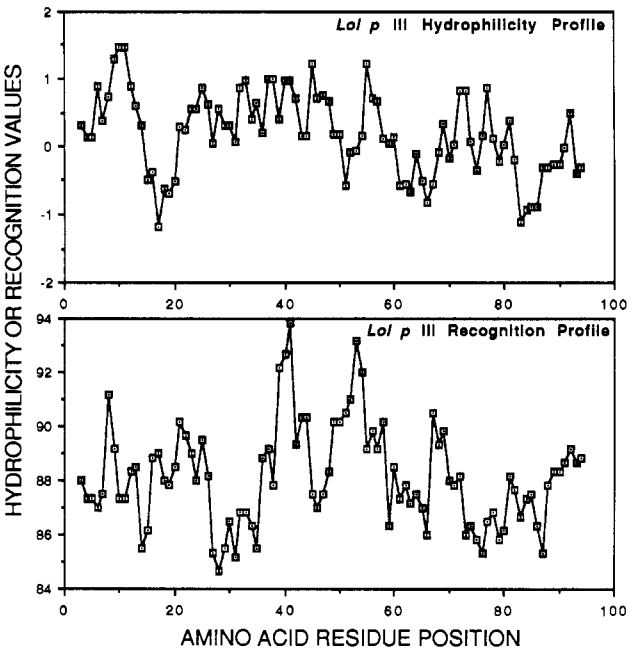

FIGURE 3: Hydrophilicity and recognition profiles for *Lol p* III. These profiles were obtained by the method of Fraga (1982) by using blocks of six residues.

Discussion

An unambiguous and complete amino acid sequence of *Lol p* III (variant IIIB) was derived from tryptic, chymotryptic, V8 protease, endoproteinase ASP-N, clostripain, and CNBr peptides of the protein (Figure 1). Like *Lol p* II, which we sequenced earlier (Ansari et al., 1989c), *Lol p* III contains 97 amino acid residues, lacks cysteine, and shows no evidence of glycosylation. The relative molecular weight of *Lol p* III, calculated from its predominant sequence, is 10 909, which is in excellent agreement with the estimate of 11 000 by physical measurements (Johnson & Marsh, 1965a,b; Marsh, 1974, 1975). The isoelectric point, calculated from amino acid data, was 9.33, which is close to the value of 9.0 obtained by isoelectric focusing. Based on the number of acidic and basic residues in *Lol p* III, a net charge of +0.5 at pH 6.5 was calculated, assuming that histidine carries a charge of +0.5 (Offord, 1966).

Except for sporadic, short-segment similarities, no sequence similar to that reported here for *Lol p* III was found in the protein sequence database (NBRF, release no. 19.0, and Swiss Protein Database, release no. 9.0). The *Lol p* III sequence was compared with our recently published complete sequence of *Lol p* II (Ansari et al., 1989c). The two molecules were found to be highly homologous: 57 of the 97 positions are occupied by identical residues (59% identity); this proportion increases to 67% when D, K, S, and I are equated with E, R, T, and V, respectively. We also compared the *Lol p* III sequence with *Lol p* I, for which the following partial sequences are known: the NH_2-terminal 30-residue sequence of *Lol p* I (Cottam et al., 1986), and a 28-residue internal sequence (Esch & Klapper, 1989) that is known to contain an important antibody-combining region of *Lol p* I and several other group I species (homologues of *Lol p* I). The NH_2-terminal sequence of *Lol p* I showed no similarity with the *Lol p* III sequence, but the antibody-combining fragment of *Lol p* I showed a high degree of homology with a portion of the *Lol p* III sequence (Figure 2).

Two theoretical analyses were performed on the *Lol p* III sequence in order to predict its potential B-cell epitopes (Hopp & Woods, 1981; Fraga, 1982; Ansari, 1985) and Ia/T-cell epitopes (Margalit et al., 1987). It should be emphasized that these are only predictive approaches, the outcome of which must be tested experimentally. Figure 3 shows the Fraga hydrophilicity and recognition profiles, the latter being a measure of interactions between neighboring amino acid residues. Following the procedure described elsewhere (Ansari, 1985), four segments that may contain or contribute to antibody-combining sites or B-cell epitopes were predicted, as

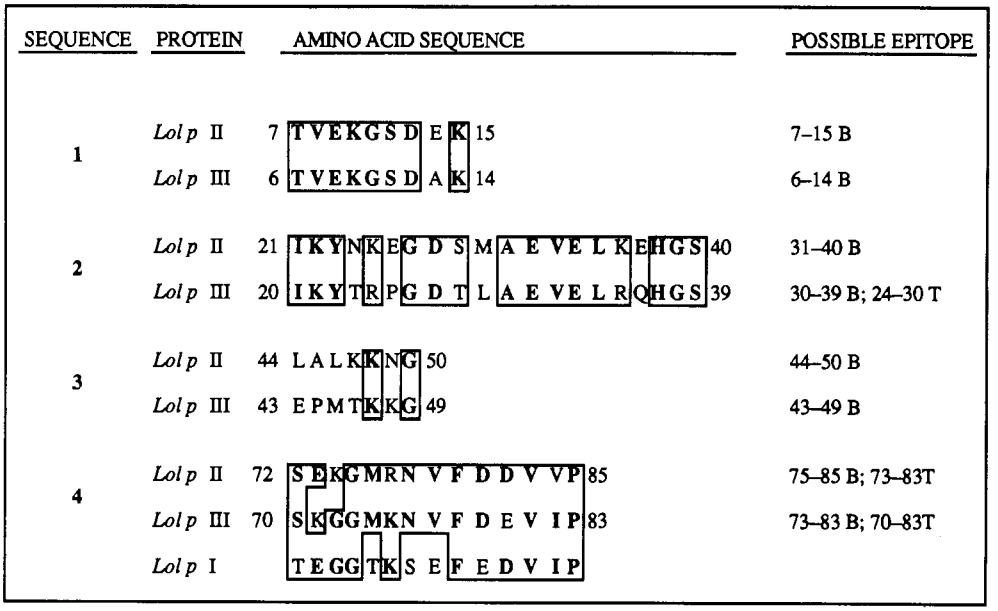

FIGURE 4: Predicted B-cell and Ia/T-cell epitopes of *Lol p* II and III. The predicted *Lol p* II epitopes are from Ansari et al. (1989c). A segment from known partial sequence of *Lol p* I (Esch & Klapper, 1989) is included for comparison because of the noted homology. Similarities between the two proteins are indicated by boxes (see legend to Figure 2). The numbers indicate the starting and ending residue positions. The letter B signifies a possible B-cell epitope predicted by the methods of Fraga (1982) and Hopp and Woods (1981), and the letter T signifies a possible Ia/T-cell epitope predicted by the amphipathicity analysis of Margalit et al. (1987).

shown in Figure 4. These may be sequential or parts of conformational B-cell epitopes. Amphipathicity analysis was performed by the method of Margalit et al. (1987) by using blocks of either 7 residues or 11 residues. Only those segments that showed amphipathicity in both the analyses and were at least 5 residues long were considered to contain potential Ia/T-cell sites. For *Lol p* III, the analysis using 7-residue blocks showed two amphipathic regions, 24–30 [amphipathic score (AS) = 18.5] and 72–83 (AS = 23.1). Both these regions were confirmed to be amphipathic in the analysis using 11-residue blocks (residues 24–28, AS = 11.9; residues 70–83, AS = 31.1). These amphipathic regions, which are potential Ia/T-cell sites, are shown in Figure 4.

It is interesting to consider the structural relatedness between the three *Lol p* allergens in the light of immunochemistry of, and genetics of human immune responsiveness to, these molecules. We reported earlier (Ansari et al., 1987) that animal antisera against the *L. perenne* allergens, or human sera from rye-allergic patients, do not exhibit any cross-reactivity either between *Lol p* I and II or between *Lol p* I and III. While many animal and human sera show high degree of cross-reactivity between *Lol p* II and III, a small proportion of patients produce antibodies specific either to *Lol p* II or to *Lol p* III. These data indicate the presence of unique B-cell epitopes on *Lol p* I–III and some B-cell epitope(s) common to *Lol p* II and III. We predicted four segments that may contain or contribute to B-cell epitopes on *Lol p* II and III (Figure 4). Three of these peptide segments, 1, 2 and 4, are almost identical in the two proteins: segment 1 shows a Glu → Ala substitution, segment 2 has a Lys → Arg and a Glu → Gln substitution, and segment 4 has a Val → Ile substitution. Segments 1, 2, and 4 may provide the antibody-cross-reactive determinants of *Lol p* II and III (also see below about segment 4). Segments 3 are very different in the two molecules and might constitute, in the respective proteins, the unique B-cell determinants, which are recognized exclusively by small proportions of allergic subjects. These theoretical predictions obviously need to be tested experimentally. In the case of *Lol p* I, segment 4 is from a 28-residue peptide that has been experimentally shown to contain a B-cell epitope. Although this segment of *Lol p* I bears sufficient similarities with *Lol p* II and III sequences to expect cross-reactivity among the three molecules, we have never detected any cross-reactivity between *Lol p* I and II, or between *Lol p* I and III. This is possible considering that single amino acid differences between homologous protein antigens can sometimes result in the production of monospecific antibodies.

With regard to the genetic control of human immune responsiveness to *Lol p* I–III, we determined that IgE and IgG Ab responsiveness to the three allergens is associated with HLA-DR3 (Freidhoff et al., 1988; Ansari et al., 1989a,b). This indicates that similar, but not necessarily identical, Ia/T-cell sites in the three allergen molecules are presented to the respective T cells by DR3 molecules or DQw2.3 molecules, the genes for which are in linkage disequilibrium with DR3 genes in the Caucasoids. Analysis of the complete protein sequences showed that segments 4 in *Lol p* II and III are highly amphipathic. According to this analysis, amphipathic regions are considered to contain Ia/T cell binding sites. By applying this algorithm, 34 of 48 (70%) known Ia/T-cell sites, and 7 of 7 known class I/T-cell sites, recognized by mice were predicted correctly (Cornette et al., 1989). Thus, it is likely that segment 4 in the three molecules may represent the common Ia/T-cell epitope which is presented by a DR3 molecule or a DQw2.3 molecule.

In conclusion, we have identified a sequence (segment 4) that is similar in *Lol p* I, II, and III (Figure 4). This segment in the respective molecules appears to contain, or contribute to, a B-cell site and an Ia/T-cell site. While the B-cell site contributed by this segment of *Lol p* I appears to be non-cross-reactive with *Lol p* II and III, the B-cell sites contributed by this segment of *Lol p* II and III may or may not be cross-reactive with each other. It will be of great interest to examine experimentally whether the postulated Ia/T-cell sites contributed by this segment in all the three molecules interact with a DR3 or a DQw2.3 molecule and account for the observed DR3 association of the immune responsiveness to the three allergens.

ACKNOWLEDGMENTS

We are grateful to Dr. E. Appella for the enzyme endoproteinase ASP-N.

SUPPLEMENTARY MATERIAL AVAILABLE

Amino acid sequence analysis yield data on *Lol p* III (Figure S-1) and HPLC profiles of various digests of *Lol p* III (Figure S-2) (3 pages). Ordering information is given on any current masthead page.

REFERENCES

Ansari, A. A. (1985) *J. Biosci. 7*, 315–322.
Ansari, A. A., Carta-Sorcini, M., Mage, R. G., & Appella, E. (1976) *J. Biol. Chem. 251*, 6798–6806.
Ansari, A. A., Kihara, T. K., & Marsh, D. G. (1987) *J. Immunol. 139*, 4034–4041.
Ansari, A. A., Marsh, D. G., & Freidhoff, L. R. (1989a) *Int. Arch. Allergy Appl. Immunol. 88*, 164–169.
Ansari, A. A., Freidhoff, L. R., Meyers, D. A., Bias, W. B., & Marsh, D. G. (1989b) *Hum. Immunol. 25*, 59–71.
Ansari, A. A., Shenbagamurthi, P., & Marsh, D. G. (1989c) *J. Biol. Chem. 264*, 11181–11185.
Cornette, J., L., Margalit, H., DeLisi, C., & Berzofsky, J. A. (1989) *Methods Enzymol.* (in press).
Cottam, G. P., Moran, D. M., & Standring, R. (1986) *Biochem. J. 234*, 305–310.
Esch, R. E., & Klapper, D. G. (1989) *Mol. Immunol. 26*, 557–561.
Fraga, S. (1982) *Can. J. Chem. 60*, 2606–2610.
Freidhoff, L. R., Kautzky, E. E., Grant, J. H., Meyers, D. A., & Marsh, D. G. (1986) *J. Allergy Clin. Immunol. 78*, 1190–1201.
Freidhoff, L. R., Kautzky, E. E., Meyers, D. A., Ansari, A. A., Bias, W. B., & Marsh, D. G. (1988) *Tissue Antigens 31*, 211–219.
Hopp, T. P., & Woods, K. R. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3824–3828.
Johnson, P., & Marsh, D. G. (1965a) *Eur. Polym. J. 1*, 63–77.
Johnson, P., & Marsh, D. G. (1965b) *Nature 206*, 935–937.
Margalit, H., Spouge, J. L., Cornette, J. L., Cease, K. B., DeLisi, C., & Berzofsky, J. A. (1987) *J. Immunol. 138*, 2213–2229.
Marsh, D. G. (1974) in *Allergology: Proceedings of the 8th Congress of the International Association in Allergology* (Yamamura, Y., Frick, O. L., Horuchi, Y., Miyamoto, S., Naranjo, P., & de Weck, A. L., Eds.) pp 381–393, Excerpta Medica, Amsterdam.
Marsh, D. G. (1975) *Antigens 3*, 271–359.
Marsh, D. G., Goodfriend, L., King, T. P., Løwenstein, H., & Platts-Mills, T. A. E. (1986) *Bull. WHO 64*, 767–770. Reprinted in (1987) *Mol. Immunol. 24*, 1383–1388.
Offord, R. E. (1966) *Nature 211*, 591–593.
Schwartz, R. H. (1985) *Annu. Rev. Immunol 3*, 237–261.

CORRECTION

Evidence for Targeted Gene Delivery to Hep G2 Hepatoma Cells in Vitro, by George Y. Wu* and Catherine H. Wu, Volume 27, Number 3, February 9, 1988, pages 887–892.

Our paper contained some data previously published in the *Journal of Biological Chemistry*, Volume 262, April 5, 1987, pages 4429–4432. We acknowledge and regret the duplication of this material.