

cDNA Cloning and Immunological Characterization of the Rye Grass Allergen *Lol p I**

(Received for publication, March 1, 1990)

Mary Perez, Glenn Y. Ishioka, Leslie E. Walker, and Robert W. Chesnut

From the Cytel Corporation, La Jolla, California 92037

The complete amino acid sequence of two "isoallergenic" forms of *Lol p I*, the major rye grass (*Lolium perenne*) pollen allergen, was deduced from cDNA sequence analysis. cDNA clones isolated from a *Lolium perenne* pollen library contained an open reading frame coding for a 240-amino acid protein. Comparison of the nucleotide and deduced amino acid sequence of two of these clones revealed four changes at the amino acid level and numerous nucleotide differences. Both clones contained one possible asparagine-linked glycosylation site. Northern blot analysis shows one RNA species of 1.2 kilobases. Based on the complete amino acid sequence of *Lol p I*, overlapping peptides covering the entire molecule were synthesized. Utilizing these peptides we have identified a determinant within the *Lol p I* molecule that is recognized by human leukocyte antigen class II-restricted T cells obtained from persons allergic to rye grass pollen.

One of the major species responsible for grass pollen allergy worldwide is the perennial rye grass *Lolium perenne*. At least three allergens found in the pollen of *L. perenne* have been isolated and purified biochemically. These allergens have been referred to as *Lol p I* ($M_r = 27,000-31,000$), *Lol p II* ($M_r = 11,000$), and *Lol p III* ($M_r = 11,000$) (1, 2). *Lol p I* appears to be the major allergen responsible for sensitivity to rye grass pollen since approximately 90% of rye grass allergic patients demonstrate skin test reactivity to purified *Lol p I* (3). Despite its immunological importance in rye grass pollen allergy, the complete primary structure of *Lol p I* is not known. To date only the amino acid sequence of the first 27 residues at the NH₂ terminus of the molecule (4) and the sequence of a 28-residue tryptic fragment (5) are known. In contrast to *Lol p I*, the complete amino acid sequence of *Lol p II* and III has recently been reported (6, 7).

Due to its immunological importance, knowledge of the complete primary structure of the *Lol p I* molecule may clarify several aspects of the humoral and cellular immunity against this allergen. For example, individuals suffering from hay fever due to rye grass hypersensitivity are known to produce varying levels of IgE antibodies against each of the *Lol p* antigens. Immunochemical studies, however, indicate that antibodies to *Lol p I* do not cross-react with *Lol p II* or III (8). Comparison of the amino acid sequences of the three *Lol p* proteins should confirm the lack of sequence similarity be-

tween *Lol p I* versus *Lol p II* and III.

In addition to humoral immunity to *Lol p I*, the amino acid sequence information for this protein should also help in clarifying the status of cellular immunity in rye grass allergic individuals. It is well documented that T lymphocytes isolated from patients suffering from hypersensitivity to house dust mite, grass pollens, and other allergens can respond specifically to the purified allergen (9-13). For example, studies by Black and Marsh (9) and Lanzavecchia *et al.* (11) indicate that peripheral blood T lymphocytes from rye grass allergic individuals can proliferate *in vitro* when stimulated with *Lol p I*. Of particular interest in our laboratory is the *Lol p I* reactivity of class II major histocompatibility complex (MHC)¹-restricted T cells since this subset of lymphocytes is known to provide lymphokines which activate B cells to secrete specific IgE antibodies. Previous studies indicate that class II MHC-restricted T cells do not recognize native antigen. Instead they appear to recognize peptide fragments of the antigen which have associated with class II MHC molecules expressed on specialized cells of the immune system called antigen presenting cells (14).

Class II MHC-restricted T cells may also play a significant role in the genetic association of immune responsiveness against *Lol p I* in rye grass allergic individuals. The production of high titers of IgE anti-*Lol p I* antibodies in allergic patients has been shown to be associated with their expression of the human leukocyte antigen (HLA) DR3, an antigen encoded by the class II MHC genes (15). Whether the HLA DR3 association of rye grass allergy is due to the preferential presentation of immunodominant *Lol p I* peptide fragments to specific T cells in the context of the DR3 molecule is unclear at present. The complete amino acid sequence of *Lol p I* should allow for the identification of peptide determinants or epitopes which are recognized by HLA DR3-restricted T cells as well as those with other HLA restrictions.

The immune response to *Lol p I* may have additional complexities in that isoelectric focusing analysis of purified *Lol p I* has revealed microheterogeneities within the protein molecule. These different heterogeneous forms of *Lol p I* have been termed isoallergens (2, 4). Since the isoallergens have distinct biochemical properties, they may also have different antigenic properties. For instance, one isoallergen may be immunogenic while another may not. Without the complete amino sequence of *Lol p I*, it has not been possible to define the biochemical basis of the microheterogeneity observed with *Lol p I* or to document the role of the individual isoallergens in the allergic response to rye grass pollen.

In the present report we have taken the first steps toward answering these questions. Described are the complete amino

* 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 nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) M35686 and M35687.

¹ The abbreviations used are: MHC, major histocompatibility complex; HBSS, Hanks' balanced salt solution; HLA, human leukocyte antigen; PBL, peripheral blood lymphocytes; PHA, phytohemagglutinin; rIL-2, recombinant interleukin-2.

acid sequences of two *Lol p I* isoallergens deduced from the cDNA sequences of two separate clones. Using the amino acid sequences, we identified a peptide fragment of the *Lol p I* molecule which is recognized by *Lol p I*-specific T lymphocytes generated from rye grass allergic patients.

MATERIALS AND METHODS

Non-defatted *L. perenne* pollen was purchased from Greer Laboratories, Lenoir, NC. The PTZ18R-*Bst*XI vector was obtained from Invitrogen. Restriction enzymes were purchased from Promega or Bethesda Research Laboratories. Oligonucleotide probes and cDNA sequencing primers were synthesized by Research Genetics, Huntsville, AL. Sequenase Version 2.0 was purchased from United States Biochemical Corporation, Cleveland, OH. The RNA size markers were from Bethesda Research Laboratories. Radiolabeled nucleotides and Hybond N Nylon membranes were purchased from Amersham Corp.

For tissue culture studies, Hanks' balanced salt solution and RPMI-1640 medium were obtained from GIBCO. Pooled AB-positive human serum and human recombinant interleukin-2 were purchased from Pel-Freez Clinical Systems and Genzyme Corporation, respectively. Ficoll-Paque was from Pharmacia LKB Biotechnology Inc. and phytohemagglutinin (PHA) came from Sigma. ICN Radionucleotides was the source of [*methyl*-³H]thymidine and Terasaki trays were purchased from Robbins Scientific. Purified *Lol p I* was supplied by Biotech Research Laboratories.

Extraction and Isolation of Poly(A)⁺ RNA—Twelve g of *L. perenne* pollen were ground in liquid nitrogen using a mortar and pestle. The liquid nitrogen was evaporated and 60 ml of extraction buffer (50 mM Tris-HCl, pH 8.5, 10 mM EDTA, 200 mM NaCl) were added immediately and mixed vigorously. An equal volume of Tris-buffered phenol/chloroform (50:50) was added to the pollen extract, and the mixture was homogenized using 10 strokes with the "A" pestle of a Wheaton glass Dounce homogenizer. Following centrifugation (8,000 × *g*, 20 min), a yellow aqueous phase was isolated. Phenol remaining in the aqueous phase was removed by extracting two times with chloroform/isoamyl alcohol (24:1). Nucleic acids were precipitated from the aqueous phase by the addition of an equal volume of isopropyl alcohol followed by incubation on ice for 20 min. A white precipitate was isolated by centrifugation (8,000 × *g*, 20 min). The resulting pellets were air-dried and dissolved in 7.5 M guanidine HCl containing 10 mM Tris-HCl, pH 7.5. One molar acetic acid (0.025 volumes) was added, the solution was mixed by inversion, 0.5 volumes ice-cold ethanol was added, and the mixture was stored overnight at -20 °C. The precipitate was isolated by centrifugation (8,000 × *g*, 20 min). Pellets were resuspended in the guanidine HCl solution, and the above precipitation step was repeated. Following centrifugation, the total RNA pellets were washed with ice-cold 70% ethanol, dried, and resuspended in water. Forty µg of poly(A)⁺ RNA was isolated by passage of total RNA over oligo(dT) cellulose columns using standard procedures (16).

cDNA Library Screening—A cDNA library was constructed using 5 µg of poly(A)⁺ RNA isolated from *L. perenne* pollen. Reverse transcriptase and oligo(dT) were used to synthesize the first strand of cDNA. Double-stranded cDNA was cloned into the vector PTZ18R-*Bst*XI using standard cloning procedures as described in the Invitrogen cloning manual. Restriction enzyme analysis of 20 different clones revealed a library insert size distribution between 0.5 and 2.5 kilobases. The unamplified library was screened using a ³²P-labeled synthetic oligonucleotide probe (5'-AGC-CAC-TTG-TC[C/G]-CCG-TAC-TC-3'). This probe is complementary to amino acids 13 through 19 of the mature *Lol p I* protein (4). This probe was designed by constructing a plant codon utilization chart similar to the human protein coding sequence chart described by Lathe (17). The frequency of codon usage for six different cloned wheat proteins was analyzed by the University of Wisconsin genetics computer group sequence analysis software packages version 6.1 (18–22). It was determined that the best NH₂-terminal amino acid sequence having a restricted number of codons resided between amino acids 13 through 19. Codon frequencies within this sequence were all greater than 89%, the exception being Gly at residue 15 which was 58%. After we had designed the probe used in this study, a plant codon utilization table was published (23). Comparison of the probe that we designed for use in this study to a probe designed using this published codon frequency chart suggested the use of an identical "guessmer" as an appropriate probe (23, 24). Hybridization conditions and filter washes were as

described in Bethesda Research Laboratories Focus publication volume 9, no. 2, with the exception that Sarkosyl was substituted for sodium dodecyl sulfate. Six potential positives were isolated from screening 40,000 colonies of a total library of 150,000.

cDNA Sequence Analysis—Double-stranded DNA was prepared by *CsCl* banding and was sequenced using the dideoxy chain-termination method employing [³⁵S]dATP and Sequenase (25). Synthetic oligonucleotide primers were used to prime the chain terminating reaction. Nucleotide and amino acid sequences were analyzed using the databases GenBank release 60 and NBRF-Protein release 20.

Northern Analysis—Poly(A)⁺ RNA (5 µg) isolated from *L. perenne* pollen was separated by electrophoresis on a 1.2% formaldehyde-agarose gel, transferred to a Hybond N Nylon membrane and hybridized to the ³²P-labeled oligonucleotide probe.

Preparation of Peripheral Blood Lymphocytes—Two-hundred fifty ml of heparinized blood was drawn from rye grass allergic individuals who previously had demonstrated a 2–4+ wheal and erythema skin reaction upon intradermal injection of 0.02 ml of a 1:10,000 dilution of *L. perenne* extract. After centrifuging the blood for 15 min at 250 × *g*, the plasma was removed and the packed cells were diluted 1:1 (v/v) with Hanks' balanced salt solution (HBSS). Twenty-five ml of these cells was layered over 15 ml of Ficoll-Paque in a 50-ml polypropylene tube. The Ficoll-Paque gradients were then centrifuged at 400 × *g* for 20 min at room temperature. Peripheral blood lymphocytes (PBL) which migrated to the HBSS-Ficoll interface were collected, washed three times with HBSS, and counted. Aliquots of cells were frozen in liquid nitrogen in 10% dimethyl sulfoxide for future use.

Generation of Anti-*Lol p I* T Cells—Anti-*Lol p I*-specific T cells were raised using procedures previously described by other laboratories (26). Initially, 10⁵ PBL were stimulated with 5 µg/ml *Lol p I* in individual wells of a U-bottom microtiter plate. Medium used for culture consisted of RPMI-1640 supplemented with 5% human serum. Seven days after initiation of culture, human recombinant IL-2 (rIL-2) was added to each well at a final concentration of 5 units/ml. Thereafter, wells were fed on a 3-day cycle with medium containing 15 units/ml rIL-2.

After addition of rIL-2, T cells in most microtiter wells grew to a high density and were subsequently transferred into individual wells of a 24-well tissue culture plate in 2 ml of medium containing 15 units/ml rIL-2. When sufficient numbers of T cells were obtained they were screened for *Lol p I* reactivity by measuring their ability to proliferate in the presence of antigen. T cells (10⁴/well) were mixed in U-bottom microtiter wells with 6–8 × 10⁴ autologous PBL which had been X-irradiated with 3000 rads. Triplicate wells were stimulated with 5 µg/ml *Lol p I*, whereas control wells did not receive antigen. After 2 days of incubation at 37 °C in 5% CO₂, [*methyl*-³H]thymidine was added at a concentration of 1 µCi/well. Twenty-four h later, cells were collected onto glass fiber discs using an automated cell harvester and counted in a Beckman liquid scintillation beta counter. T cells demonstrating [*methyl*-³H]thymidine incorporation 3-fold above control wells were considered to be specific for *Lol p I* and were immediately cloned.

Anti-*Lol p I*-specific T cells were cloned at limiting dilution in individual wells of a 72-well Terasaki typing tray in 20 µl of medium containing 1 µg/ml phytohemagglutinin (PHA), 15 units/ml rIL-2, and 2 × 10⁴ allogeneic irradiated PBL. Seven to 10 days after culture, individual Terasaki wells were examined under a microscope, and wells containing actively growing cells were transferred to U-bottom wells containing rIL-2 medium. T cells were considered clonal in nature if they were derived from Terasaki trays in which less than 37% of the wells contained growing cells. Within 3 days after transfer into U-bottom wells, T cell clones were restimulated with PHA (1 µg/ml) and irradiated allogeneic PBL (10⁵/well). Cells were eventually expanded into larger wells in medium containing 15 units/ml rIL-2 before being tested for antigen specificity in a lymphocyte proliferation assay. T cell clones which showed greater than 3-fold incorporation of [*methyl*-³H]thymidine above background upon antigen stimulation were selected for further study.

T cell clones were grown routinely by restimulating them on a 14–21-day cycle with PHA. One to three × 10⁵ T cells were added to wells in 2 ml of medium containing 1 µg/ml PHA, 15 units/ml rIL-2, and 1.5 × 10⁶ irradiated allogeneic PBL. Three days after PHA stimulation, T cells were transferred to larger wells and fed on a 3-day cycle with rIL-2 medium. Cells were used for lymphocyte proliferation assays 7 days after PHA stimulation.

Identification of T Cell Determinants in the *Lol p I* Molecule—To identify T cell determinants in the *Lol p I* molecule, T cell clones were tested for their ability to proliferate when stimulated with a set

T cell proliferation was measured as described above by mixing 1.5×10^4 T cells, 8×10^4 irradiated autologous PBL, and Lol p I peptide at a final concentration of $10 \mu\text{g/ml}$ in U-bottom wells. Each overlapping peptide was tested in duplicate wells. Peptides which stimulated [*methyl*- ^3H]thymidine incorporation in T cell clones 3-fold above background were considered potential T cell determinants. Each T cell peptide determinant was retested to confirm reactivity.

Isolation and Characterization of *Lol p I* cDNA Clones—Twenty-seven of 30 NH₂-terminal amino acids of the *Lol p I* antigen have been previously identified by automated amino acid sequencing (4). A synthetic oligonucleotide probe complementary to amino acids 13 through 19 was end labeled with ³²P and used to screen a *L. perenne* pollen cDNA library in PTZ18R-*Bst*XI. Six clones were identified among 40,000 colonies which had 1.2-kilobase cDNA inserts hybridizing to the oligonucleotide probe. Restriction enzyme analysis of the six clones revealed that one of the clones, 5A, had an internal *Eco*RI site not detected in the other five clones (see Fig. 1).

The cDNA did not contain the initiator Met, but the amino acid sequence predicted by the cDNA matches the partial NH₂-terminal protein sequence previously determined, confirming the identity as authentic *Lol p I*, and that it codes for the mature protein (4). The NH₂-terminal 12 amino acids probably represent a portion of the signal peptide since the cDNA NH₂-terminal sequence is not contained in the mature protein. Furthermore, as has been noted for other signal peptides, it is hydrophobic and contains small side chain amino acids at the -1 and -3 positions relative to the signal peptide cleavage site (28).

sequence asparagine-X-threonine. Chemical methods have suggested that the *Lol p* I allergen contains approximately 4% carbohydrate (4).

Clone 5A, a separate isolate, was also sequenced to completion, and the nucleotide and deduced amino acid sequence are shown in Fig. 3. Clone 5A contains the complete proposed hydrophobic signal sequence and also codes for a 240-amino acid mature protein. Comparison of clone 1A with clone 5A reveals four amino acid differences. One results in a change in the net charge of the *Lol* p I antigen; Asp-45 of clone 1A is an Asn in clone 5A. Other nucleotide changes resulting in amino acid changes are the following: Gly-154 (clone 1A) is Ala in clone 5A, Ile-187 (clone 1A) is Thr, and Val-223 (clone 1A) is a Phe. There are 20 additional individual nucleotide differences in clone 5A as compared with clone 1A which do not result in amino acid changes (Fig. 3).

Both clones contain an extensive 3'-untranslated region. Although the 3'-untranslated regions are similar, clone 5A contains an additional 51 nucleotides.

The complete amino acid sequence of *Lol p* II has recently been reported (6). Alignment of the amino acid sequence of *Lol p* I with *Lol p* II starting at the carboxyl-terminal end of each protein and comparison of their amino acid sequence reveals 35 of the 97 amino acids of *Lol p* II are identical to *Lol p* I (Fig. 4). A computer-assisted homology search of the data bases did not reveal any homologies with other known proteins.

Comparison of clone 1A and clone 5A by Hopp and Woods (29) analysis did not show any differences in predicted protein antigenic determinants even though there were four amino acid differences between the two isoallergens.

Northern Blot Analysis—In order to characterize the mRNA for the *Lol p I* antigen, a Northern blot analysis was performed on poly(A)⁺ RNA. A single 1.2-kilobase RNA species was detected (Fig. 5) using the ³²P-labeled oligonucleotide probe. As would be expected, mammalian-derived RNA used as a control did not hybridize to the *Lol p I* probe.

Identification of T Cell Determinants in the Lol p I Allergen—In order to begin to understand the mechanisms involved in the allergic response to grass pollen antigens, an effort was made to identify T cell epitopes within the Lol p I molecule. To accomplish this, anti-Lol p I-specific T lymphocyte clones generated from a total thus far, of three rye grass allergic individuals were incubated with a series of synthetic overlapping Lol p I peptides and autologous, irradiated PBL which served as a source of antigen presenting cells. Those peptides which can stimulate T cell proliferation, as measured by tritiated thymidine incorporation, are thus identified as containing a T cell determinant.

The anti-*Lol p I* T cells used in this study were raised by stimulating PBL from rye grass allergic patients with purified *Lol p I* in individual microtiter wells. Seven days after initiation of culture, rIL-2 was added to each well to expand the number of T cells. Based on the clonal frequency of antigen-reactive T lymphocytes in the peripheral blood, a certain proportion of the microtiter wells would be expected to contain T cells specific for *Lol p I*. In representative results from one individual, when T cells were screened against *Lol p I*, 18 out of a total of 184 wells screened contained antigen-specific T cells as measured by their ability to proliferate upon antigen stimulation (data not shown). Anti-*Lol p I* T cells from three of the 18 wells were subsequently cloned at limiting dilution and four T cell clones responding to purified *Lol p I* were isolated.

The T cell determinant(s) recognized by each of the four clones was identified by assaying these cells for their proliferative response to a panel of peptides.

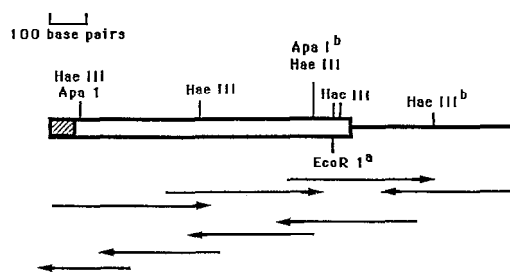
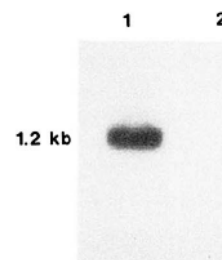
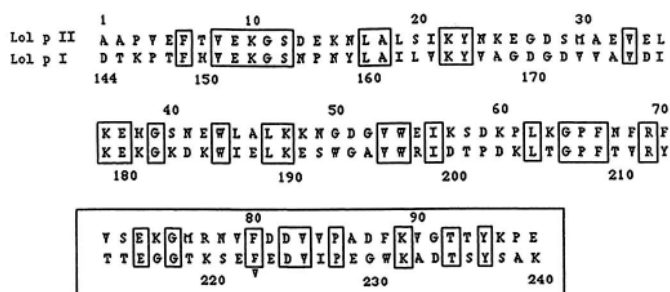


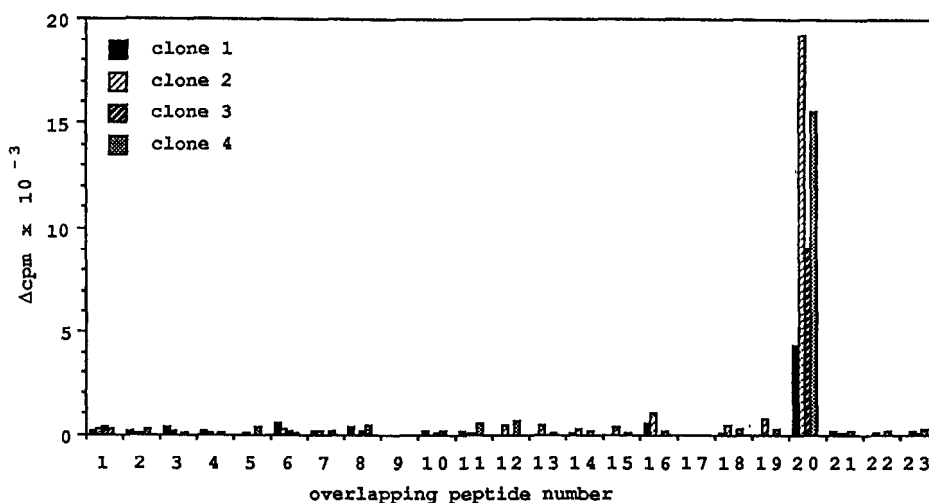
FIG. 1. Restriction map and sequencing strategy for *Lol p I*. The hatched region encodes the putative signal peptide, the open box encodes the mature protein, and the solid line indicates the 3'-untranslated region. Custom primers were used to sequence *Lol p I* and the strategy is shown by the arrows. *restriction site is in clone 5A only; †restriction site is in clone 1A only.

CGC CTG TTC GCC GTG TTC CTG GGC AGC GCG CAT GGC ATC GCG AAG GTG	48	ATG GCG TCC TCC TCG TCG GTG CTC CTG GTG GTG	33
Ala Leu Phe Ala Val Phe Leu Gly Ser Ala His Gly <u>Ile Ala Lys Val</u>	4	Met Ala Ser Ser Ser Ser Val Leu Leu Val Val	
CCC CCG GGC CCC AAC ATC ACG GCC GAG TAC GGC GAC AAG TGG CTG GAC	96	CGC CTG TTC GCC GTG TTC CTG GGC AGC GCG CAT GGC ATC GCG AAG GTA	81
Pro <u>Pro Gly</u> Pro Asn <u>Ile Thr Ala Glu Tyr Gly Asp Lys Trp Leu Asp</u>	20	Ala Leu Phe Ala Val Phe Leu Gly Ala His Gly <u>Ile Ala Lys Val</u>	4
CGC AAG AGC AAC TGG TAT GGC AAG CCG ACC GGC GCC GGT CCC AAG GAC	144	CCA CCG GGC CCC AAC ATC ACG GCC GAG TAC GGC GAC AAG TGG CTG GAC	129
<u>Ala Lys Ser Thr Trp Trp Tyr Gly Lys Pro Thr Gly</u> Ala Gly Pro Lys Asp	36	Pro Pro <u>Gly Pro</u> Asn <u>Ile Thr Ala Glu Tyr Gly Asp Lys Trp Leu Asp</u>	20
AAC GGC GGC GCG TGC GGG TAC AAG GAC GTT GAC AAG GCG CCG TTC AAC	192	GCG AAG AGC ACC TGG TAT GGC AAG CCG ACC GGC GCC GGT CCC AAG GAC	177
Asn Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys Ala Pro Phe Asn	52	<u>Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Ala Gly Pro Lys Asp</u>	36
GGC ATG ACC GGC TGC GGC AAT ACC CCC ATC TTC AAG GAC GGC CGC GGC	240	AAC GCG GGC GCG TGC GGG TAC AAG AAT GTC GAC AAG GCG CCG TTC AAC	225
Gly Met Thr Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys Asp Gly Arg Gly	68	Asn Gly Gly Ala Cys Gly Tyr Lys Asn Val Asp Lys Ala Pro Phe Asn	52
TGC GGC TCC TGC TTC GAG ATC AAG TGC ACC AAG CCC GAG TCT TGC TCC	288	GGC ATG ACC GGC TGC GGC AAC ACC CCC ATC TTC AAG GAC GGC CGT GGC	273
Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro Glu Ser Cys Ser	84	Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys Asp Gly Arg Gly	68
GGC GAG GCT GTC ACC GTC ACA ATC ACC GAC GAC AAT GAG GAG CCC ATC	336	TGC GGC TCC TGC TTC GAG ATC AAG TGC ACC AAG CCC GAG TCC TGC TCC	321
Gly Glu Ala Val Thr Val Thr Ile Thr Asp Asp Asn Glu Glu Pro Ile	100	Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro Glu Ser Cys Ser	84
GCA CCC TAC CAT TTC GAC CTC TCG GGC CAC GCG TTC GGG TCC ATG GCC	384	GGC GAG GCT GTC ACC GTC ACA ATC ACC GAC GAC AAC GAG GAG CCC ATC	369
Ala Pro Tyr His Phe Asn Leu Ser Gly His Ala Phe Gly Ser Met Ala	116	Gly Glu Ala Val Thr Val Thr Ile Thr Asp Asp Asn Glu Glu Pro Ile	100
AAG AAG GGC GAG GAG CAG AAG GTC CCG AGC GCC GGC GAG CTG GAG CTC	432	GCA CCC TAC CAT TTC GAC CTC TCG GGC CAC GCG TTC GGG TCC ATG GCC	417
Lys Lys Gly Glu Glu Gln Asn Val Arg Ser Ala Gly Glu Leu Glu Leu	132	Ala Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe Gly Ser Met Ala	116
CAG TTC AGG CGG GTC AAG TGC AAG TAC CCG GAT GAC ACC AAG CCA ACG	480	AAG AAG GGC GAG GAG CAG AAC GTC CGC AGC GCC GGC GAG CTG GAG CTC	465
Gln Phe Arg Arg Val Lys Cys Lys Tyr Pro Asp Asp Thr Lys Pro Thr	148	Lys Lys Gly Glu Glu Gln Asn Val Arg Ser Ala Gly Glu Leu Glu Leu	132
TTC CAC GTC GAG AAG GGT TCC AAC CCC AAC TAC CTG GCT ATT CTG GTG	528	CAG TTC AGG CGG GTC AAG TGC AAG TAC CCG GAC GAC ACC AAG CCG ACA	513
Phe His Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val	164	Gln Thr Arg Arg Val Lys Cys Lys Tyr Pro Asp Asp Thr Lys Pro Thr	148
AAG TAC GTC GAG GGC GAC GGC GAC GTG GTC GGC GAC ATC AAG GAG	576	TTC CAC GTC GAG AAG GCT TCC AAC CCC AAC TAC CTC GCT ATT CTG GTG	561
Lys Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val Asp Ile Lys Glu	180	Phe His Val Glu Lys Ala Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val	164
AAG GGC AAG GAT AAG TGG ATC GAG CTC AAG GAG TCG TGG GGA GCA GTC	624	AAG TAC GTC GAC GGT GAC GGT GAC GTG GTG GCG GTG GAC ATA AAG GAG	609
Lys Gly Lys Asp Lys Trp Ile Glu Leu Lys Glu Ser Trp Gly Ala Val	196	Lys Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val Asp Ile Lys Glu	180
TGG AGG ATC GAC ACC CCC GAT AAG CTG ACG GGC CCA TTC ACC GTC CGC	672	AAG GGC AAG GAC AAG TGG ACC GAG CTT AAG GAG TCG TGG GGA GCA GTC	657
Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro Phe Thr Val Arg	212	Lys Gly Lys Asp Lys Trp Thr Glu Leu Lys Glu Ser Trp Gly Ala Val	196
TAC ACC ACC GAG GGC GGC ACC AAA TCC GAA GTC GAG GAT GTC ATT CCC	720	TGG AGG ATC GAC ACC CCC GAT AAG CTG ACC GGC CCA TTC ACC GTC CGC	705
<u>Tyr Thr Thr Glu Gly Thr Lys Ser Glu Val Glu Asp Val Ile Pro</u>	228	Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro Phe Thr Val Arg	212
GAG GGC TGG AAG GCC GAC ACT TCC TAC TCG GCC AAG TGA	759	TAC ACC ACC GAG GGC GGC ACC AAA TCC GAA TTC GAG GAT GTC ATC CCT	753
<u>Glu Gly Trp Lys Ala Asp Thr Ser Tyr Ser Ala Lys</u>	240	<u>Tyr Thr Thr Glu Gly Thr Lys Ser Glu Phe Glu Asp Val Ile Pro</u>	228
GCAAGAGTGAAGTGATCTCTTCGACGAGCTTAATTTTGACTCAAGATCTCAATAATCCAG	822	GAG GGC TGG AAG GCC GAC ACC TCC TAC TCG GCC AAG TGA	798
CCACACATATATACGAGCGGTGACATACAAAGCTATGCATGAGTATATTCATTACGCCG	888	<u>Glu Lys Thr Lys Ala Asp Thr Ser Tyr Ser Ala Lys</u>	240
TATATAGAGAGGAGAGATGCGTGAATAAGAGTTTGATGCTGACACCTTGTGAGAGGTGTG	954	CGAGGAAGTGAAGTGATCTTCGTCCAATCAGCTTAATTTTGACTCAAGATCTCAATAATCTAG	861
AGGAGGCACCCAATCAAAATTTGGTTGGCCCTCCCGCCCCAGCGTGAGCTGACTGATTATGAT	1020	CGCAGCATCATATCGAGGGGTGAGACATACAACTCTCATGAATGAGTATATTCATTCATG	924
CTAAGACTGTTTCATGATTACTATTAATAAAAAAAAAAAAAAAAAAAAAA	1066	CCGATATGAGAGGAGAGGTAGCTGACCAATAGTTTGGAGGTCTGACACCTTGTGAGGTGGAGA	987
		GGTCAGAGGTGTAAAGTACGGGAGCACTAATCAAAATTTGAGAGCCGCCCCCTCTAGAGT	1050
		CTTATTCATGTACCTAATCTGTTTCATGATTCTGAAACACATTAAATGATTTTTTACACATT	1113
		AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1150
		3'	



6, when the four T cell clones (numbered 1 through 4) were tested against each of the peptides in the presence of irradiated autologous PBL serving as antigen presenting cells, only peptide number 20 induced significant proliferation from all of the clones. Thus, peptide number 20, containing the amino acid residues 191-210 (ESWGAVWRIDTPDKLTGPFT) (see Figs. 2 and 3), appears to contain an immunodominant

FIG. 6. Identification of T cell determinants within the *Lol p I* molecule. Four anti-*Lol p I* T cell clones, designated as 1–4, were tested for their ability to proliferate in response to 23 overlapping synthetic peptides covering the entire *Lol p I* molecule (see "Materials and Methods" for a description of the assay). Results are expressed as Δ cpm where the incorporation of [*methyl*]-³H]thymidine (in cpm) into DNA by the T cells in the presence of peptide was subtracted by the incorporation in the absence of peptide. Proliferation of T cells in the absence of peptide ranged from 185 to 982 cpm.



determinant recognized by T cells from this rye grass allergic individual. T cell clones isolated from two additional rye grass allergic patients also responded to peptide number 20. Additional clones are now being generated and tested in a similar manner.

DISCUSSION

The complete amino acid sequence of two *Lol p I* "isoallergens" have been deduced from the cDNA sequence of two separate clones. Comparison of the deduced amino acid sequence of clone 1A (Fig. 2) with clone 5A (Fig. 3) reveals 25 nucleotide differences. However, these differences result in only four amino acid changes. The most notable change is Asp-45 (clone 1A) which is Asn-45 (clone 5A), thus resulting in a difference in charge. This charge difference may account for the *Lol p I* microheterogeneity observed following analytical isoelectric focusing (4). In addition, the amino acid sequence predicted by the cDNA sequence correlates with the published amino acid composition analysis (4, 30).

Partial cDNA sequence analysis of four additional cDNA clones (data not shown), showed sequence identity to clone 1A (Fig. 2). These results suggest, although do not prove, that the amino acid sequence deduced from clone 1A represents the predominant form of the allergen. In addition, although the function of *Lol p I* in rye grass pollen is not known, the relatively low number of amino acid changes (4 out of 270) may suggest some evolutionary pressure to maintain *Lol p I* structure. However, the limited number and conservative nature of the amino acid substitutions may reflect the limited number of clones characterized in this analysis, the specific pollen batch used to make the RNA, or the specificity of the oligonucleotide probe used to select the *Lol p I* clones. This question can only be addressed as additional cDNAs are cloned and sequenced.

Since the *Lol p I* protein is expressed as isoallergenic forms, the question arises whether the forms identified are able to elicit an allergic response. To address this question, experiments were designed to determine whether the two cDNA clones isolated and sequenced could code for biologically active rye grass allergens and, if so, to try to identify a peptide moiety(s) involved in the allergic response. For these experiments, T cell clones against purified *Lol p I* were generated from the peripheral blood cells of rye grass allergic individuals. These clones were subsequently screened for their reactivity against a set of overlapping peptides, the sequence of which was derived from the cDNA clones and which covered the entire *Lol p I* molecule. T cell clones were identified from

each of the individuals studied that responded to a 20-amino acid synthetic peptide containing residues 191–210 of the *Lol p I* molecule. Whether this peptide represents an immunodominant T cell determinant will require examination of a greater number of T cell clones and a larger population of patients. Work is in progress to accomplish this. We attempted to determine the class II restriction element of one of the T cell clones which responded to the *Lol p I* 191–210 peptide. Using fibroblasts transfected with HLA class II genes as the antigen presenting cell, clone 4 proliferated in response to peptide presented by fibroblasts expressing HLA DRw52a consistent with the HLA haplotype of the patient from which clone 4 was derived (data not shown). Thus, T cell recognition of the *Lol p I* 191–210 peptide occurs in the context of the HLA DRw52a molecule.

Studies are currently in progress to identify epitopes recognized by T cells generated from additional rye grass-allergic individuals who express different HLA class II haplotypes including HLA DR3. Any role of T cells in the DR3 association of allergy to *Lol p I* will be revealed as more antigenic determinants are identified and their HLA restriction defined. Although we have determined that recognition of the *Lol p I* 191–210 peptide is restricted by HLA DRw52a, Ansari *et al.* (6, 7) have recently proposed that another T cell determinant restricted by HLA DR3 may exist within the tryptic fragment previously isolated by Esch and Klapper (5). This prediction was based on the algorithm deduced by DeLisi and Berzofsky (31) which indicated that there was a high propensity for amphipathic α -helical peptides to form T cell epitopes. We now know that this tryptic fragment, containing residues 213–240, is located at the carboxyl terminus of the *Lol p I* molecule. The immunogenic nature of this carboxyl-terminal peptide should be revealed by analyzing the specificity of T cell clones derived from allergic DR3 individuals.

Another interesting aspect of the *Lol p I* tryptic fragment isolated by Esch and Klapper is its high similarity to residues 72–97 in the carboxyl terminus of *Lol p II* (see Fig. 4), as reported by both these investigators and by Marsh *et al.* (6, 7). Comparison of the complete amino acid sequence of *Lol p I* with *Lol p II* reveals that 34 of the 97 amino acid residues of *Lol p II* are contained within residues 144 through 240 of *Lol p I*. In addition to the sequence homology between *Lol p I* and *II*, other biochemical studies suggest the *Lol p I* may have similar structure to pollen allergens isolated from related grasses (5, 32). For example, polyclonal and monoclonal antibodies raised against *Lol p I* show cross-reactivity to allergens of similar molecular weight isolated from the pollen of

orchard, fescue, kentucky blue grass, timothy grass, and several other grass species (32–34). Comparison of the amino acid sequence of these related pollen allergens, when they become available, may provide further insight into the observed cross-reactivity with specific antibody. Furthermore by using anti-*Lol p I* T cell clones, studies can be undertaken to address the cross-reactivity between *Lol p I* and II and between *Lol p I* and purified allergens from related grasses at the T cell level.

Acknowledgment—We would like to thank Donald Stevenson M.D., Department of Clinical Immunology, Scripps Clinic and Research Foundation, La Jolla, California for access to blood samples from rye grass-sensitive patients.

REFERENCES

- Johnson, P., and Marsh, D. G. (1965) *Eur. Polym. J.* **1**, 63–77
- Johnson, P., and Marsh, D. G. (1965) *Nature* **206**, 935–937
- Friedhoff, L. R., Ehrlich-Kautzky, E., Grant, J. H., Meyers, D. A., and Marsh, D. G. (1986) *J. Allergy Clin. Immunol.* **78**, 1190–1201
- Cottam, G. P., Moran, D. M., and Standring, R. (1986) *Biochem. J.* **234**, 305–310
- Esch, R. E., and Klapper, D. G. (1989) *Mol. Immunol.* **26**, 557–561
- Ansari, A. A., Shenbagamurthi, P., and Marsh, D. G. (1989) *J. Biol. Chem.* **264**, 11181–11185
- Ansari, A. A., Shenbagamurthi, P., and Marsh, D. G. (1989) *Biochemistry* **28**, 8665–8670
- Ansari, A. A., Kihara, T. K., and Marsh, D. G. (1987) *J. Immunol.* **139**, 4034–4041
- Black, P. L., and Marsh, D. G. (1980) *J. Allergy Clin. Immunol.* **66**, 394–401
- Buckley, R. H., Seymour, F., Sanal, S. O., Ownby, D. R., and Becker, W. G. (1977) *J. Allergy Clin. Immunol.* **59**, 70–78
- Lanzavecchia, A., Santini, P., Maggie, E., Del Prete, G. F., Falagiani, P., Romagnani, S., and Ferrarini, M. (1983) *Clin. Exp. Immunol.* **52**, 21–28
- O'Hehir, R. E., Young, D. B., Kay, A. B., and Lamb, J. R. (1987) *Immunology* **62**, 635–640
- Rawle, F. C., Mitchell, E. B., and Platts-Mills, T. A. E. (1984) *J. Immunol.* **133**, 195–201
- Shimonkevitz, R., Kappler, J., Marrack, P., and Grey, H. (1984) *J. Exp. Med.* **158**, 303–316
- Friedhoff, L. R., Ehrlich-Kautzky, E., Meyers, D. A., Ansari, A. A., Bias, W. B., and Marsh, D. G. (1988) *Tissue Antigens* **31**, 211–219
- Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Lathe, R. (1985) *J. Mol. Biol.* **183**, 1–12
- Baulcombe, D. C., Huttly, A. K., Martienssen, R. A., Barker, R. F., and Jarvis, M. G. (1987) *Mol. Gen. Genet.* **209**, 33–40
- Forde, J., Forde, B. G., Fry, R. P., Kreis, M., Shewry, P. R., and Mifflin, B. J. (1983) *FEBS Lett.* **162**, 360–366
- Litts, J. C., Colewell, G. W., Chakerian, R. L., and Quantrano, R. S. (1987) *Nucleic Acids Res.* **15**, 3607–3618
- Smith, S. M., Bedbrook, J., and Speirs, J. (1983) *Nucleic Acids Res.* **11**, 8719–8734
- Tabata, T., Sasaki, K., and Iwabuchi, M. (1983) *Nucleic Acids Res.* **11**, 5865–5875
- Murray, E. E., Lotzer, J., and Eberle, M. (1989) *Nucleic Acids Res.* **17**, 477–498
- Maniatis, T., Fritsch, E. F., and Sambrook, J. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Sanger, F., Nicklen, S., and Coulson, A. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
- Gaston, J. S. H., Life, P. F., Jenner, P. J., Colston, M. J., and Bacon, P. A. (1990) *J. Exp. Med.* **171**, 831–841
- Sette, A., Buus, S., Colon, S., Miles, C., and Grey, H. M. (1988) *J. Immunol.* **141**, 45–48
- Kozak, M. (1987) *Nucleic Acids Res.* **15**, 8125–8148
- Hopp, T. P., and Woods, K. R. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 3824–3828
- Johnson, P., and Marsh, D. G. (1966) *Immunochemistry* **3**, 91–100
- DeLisi, C., and Berzofsky, J. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 7048–7052
- Marsh, D. G., Haddad, Z. H., and Campbell, D. H. (1970) *J. Allergy* **46**, 107–121
- Kahn, C. R., and Marsh, D. G. (1986) *Mol. Immunol.* **23**, 1281–1288
- Esch, R. E., and Klapper, D. G. (1987) *J. Allergy Clin. Immunol.* **79**, 489–495