Tertiary Structure of the Major House Dust Mite Allergen Der p 2: Sequential and Structural Homologies[†,‡]

Geoffrey A. Mueller,[§,||] David C. Benjamin,[§] and Gordon S. Rule*

Beirne B. Carter Center for Immunology, the Asthma and Allergic Disease Center, and the Biophysics Program, University of Virginia, and Carnegie Mellon University, Department of Biological Sciences, 4400 Fifth Avenue, Pittsburgh, Pennsylvania 15213

Received March 13, 1998; Revised Manuscript Received June 10, 1998

ABSTRACT: Sensitization to indoor allergens, especially those of the house dust mite, is strongly correlated with the development of asthma. We report the tertiary structure of the major house dust mite allergen, Der p 2, determined by NMR methods. The structure of Der p 2 is a β-barrel and is composed of two three-stranded antiparallel β-pleated sheets. This arrangement of β-strands is similar to the immunoglobulin fold with respect to the orientation of the two sheets and the interactions of the strands. However, the three-dimensional structure of Der p 2 aligns equivalently with a number of proteins from different families within the immunoglobulin superfamily. The structural homology with the highest significance score from analysis by DALI is to Der f 2. Although Der p 2 and Der f 2 are 87% identical in amino acid sequence, they align in three dimensions rather poorly (4.85 Å RMSD; Z-score, 8.58). This unexpected finding is likely due to the different solution conditions used during structure determination by NMR for both proteins. While the structural comparisons did not elucidate a clear homologue for the function of Der p 2 in mites, we report that Der p 2 is sequentially homologous to esr16. This is a protein from moths that is expressed coincident with molting. Thus, this homology has important ramifications for the study of mite allergy. The structure of Der p 2 provides a useful tool in the design of recombinant immunotherapeutics for the group 2 allergens.

Throughout most of the world, the most common inhalant allergens associated with asthma are those from the house dust mite (*1*). Up to 80% of asthmatic children are sensitized to the house dust mite (*2*). Among the mite allergens, Der p 2 and the homologous Der f 2 (87% sequence identity to Der p 2) are considered major allergens because greater than 85% of patients who react to mite extract react specifically with these group 2 allergens (*1*, *3*).

Der p 2 does not fit the typical paradigm for dust mite allergens. Most allergens from the house dust mite are found in the fecal particles, but Der p 2 occurs primarily in whole body extract rather than spent waste (*4*). Many of the dust mite allergens are lytic enzymes (*5*), suggesting that the function of an allergen could be related to its pathogenicity (*6*). Der p 2 does not have a homologue with known enzymatic activity and the identification of such a homologue would aid in the understanding of the epidemiology of Der p 2. Prior to this study, Der p 2 was only known to have a 35% sequence identity with a human epidydimal gene product (HE1), suggesting an undefined role of Der p 2 in mite reproduction (*7*). The structure of Der p 2 reported here facilitates the identification of additional proteins and enzymes which may show structural homology to Der p 2.

The availability of recombinant allergens has facilitated the use of site-directed mutagenesis to create modified antigens for immunotherapy. This approach to develop an improved recombinant allergen is being explored for the indoor allergens Der p 2 and Der f 2 as well for the birch pollen allergen Bet v 1 and the peanut allergens Ara h 1 and Ara h 2 (*8*−*10*). During immunotherapy, an allergic patient is repeatedly injected with an allergen with the goal of reducing symptoms upon subsequent contact with the allergen in a natural setting. However, a significant risk of systemic anaphylaxis accompanies any allergen-specific desensitization protocol due to reaction of the injected allergen with preexisting IgE on the surface of mast cells and basophils (*11*). An ideal recombinant allergen would have a reduced risk of anaphylaxis, yet maintain its ability to desensitize.

To efficiently develop modified antigens for immunotherapy, it is necessary to determine the location of epitopes on the surface of the allergen. Epitopes have been mapped for the allergens Bet v 1, Ara h 1, and Ara h 2 using peptides which are presumably in a nonnative configuration (*8*−*10*). In contrast, peptides of Der p 2 show very little reactivity with patient IgE, implying that the epitopes are dependent on the tertiary structure of the protein (*12*−*14*). The three disulfide bonds in Der p 2 and Der f 2 [residues 21−27, 73−78, and 8−119 (*15*)] appear to be important for recognition by antibody. B-cell epitopes of the group 2

[†] Supported by a grant from the NIH (AI-34607) and the Eberly Family Professorship in Structural Biology (G.S.R.).

[‡] Coordinates have been deposited at the Brookhaven Protein Data Bank, accession code is 1A9V.

* Corresponding author at Carnegie Mellon University. Phone: (412)-268-1839. E-mail: rule@andrew.cmu.edu.

[§] Beirne B. Carter Cener for Immunology and the Asthma and Allergic Disease Center.

[||] Biophysics Program.

S0006-2960(98)00578-9 CCC: $15.00
Published on Web 08/22/1998

allergens were shown to be sensitive to reduction or alkylation (*16*). The importance of the disulfides was further supported by alanine-scanning mutagenesis of Der f 2, which showed that disruption of any of the three disulfide bonds had significant effects on IgE binding (*17*). More recently, mutants of Der p 2 and Der f 2 which lack one or more of the three disulfides have been examined for their ability to elicit reduced skin test responses in mite allergic individuals (*18*, *19*).

Additional studies have investigated the role of other residues in Der p2 and Der f2 in antibody binding. Smith and Chapman used predictive algorithms to identify surface residues that may be involved in the interaction between Der p 2 and antibody and suggest that residues 44−46 and 100 are contained within epitopes recognized by murine and human antibodies (*20*). A related study on Der f 2 identified residues 15, 19, 71−73, 76−78, and 119−121 as being important in antibody binding (*17*).

The above studies have made significant contributions to the understanding of the immunology of group 2 allergens. However, it is difficult to interpret these results or that from any other modified allergen without knowledge of the tertiary structure of the protein and its immune complexes. Given the conformational nature of the antibody epitope in Der p2 and Der f2, detailed structural information will be of high value in the design of recombinant material for immunotherapy.

The NMR structure of Der f 2 has been recently reported (*21*). This structure contains 8 β-strands that form a pattern typical of the Immunoglobulin fold. Der f 2 is reported to contain 47% sheet and 53% random coil based on chemical shift index prediction and an examination of the long-range NOEs. Preliminary NMR studies on Der p 2 using similar criteria suggested a secondary structure of 49% sheet and 51% random coil (*22*). Although the secondary structure of Der f 2 and Der p 2 are similar, there are significant differences in the overall tertiary fold despite the high-sequence identity. The structural differences between these two allergens will be discussed in detail, and we conclude that these differences can be accounted for by the very different solution conditions employed for NMR spectroscopy. The structure of Der p 2 reported here was obtained under more physiological conditions. It clarifies much of the existing data regarding the immunology of Der p 2 and its variants and provides a foundation for the design of new variants of Der p 2 for immunotherapy.

EXPERIMENTAL PROCEDURES

Expression of Der p 2. The recombinant protein used in this study was previously characterized to be similar to the natural mite product by competitive inhibition ELISA with a panel of mAb, and a radio-allergosorbent test (RAST)[1] with patient sera (*22*).

Structure Determination. The experiments used to determine the NMR structure of Der p 2 were previously described (*22*). All spectra of rDer p 2 (D1S) for assignment and structure determination purposes were obtained at 25 °C in a buffer of 10 mM sodium phosphate, pH 6.0, 50 mM NaCl, 100 mM K_2SO_4, 5 μM EDTA, sodium azide (0.02% w/v), and D_2O (5% v/v) (P2 buffer). To examine the solution conditions under which the Der f 2 structure was determined, an $^{15}N/^{1}H$ HSQC (*23*) was obtained with Der p 2 in 140 mM *N*-octyl-β-D-glucoside and 10% D_2O, (F2 buffer) (*21*).

Virtually complete resonance assignments of the aliphatic and amide resonances were made using standard triple resonance techniques (*22*). A number of additional aromatic resonances were assigned with a 2D homonuclear NOE SCUBA (*24*) and a $^{13}C/^{1}H$ HSQC. All Hδ and Hϵ protons on the three Tyr and two Phe residues were assigned. Three of the five His Hδ protons were assigned. Assignments for the HE3, HD1, and HE1 of the single Trp residue (residue 92) were also obtained.

NOE restraints were determined in three phases. During the first phase, an NOE restraint was included in structure modeling only if both the proton and the corresponding heteronuclear chemical shift uniquely identified a single possible restraint. For example, the cross-peak in the 3D-CCH-NOESY spectrum had to have a corresponding cross-peak in the 3D-HCH-NOESY spectrum (*25*). The second phase was initiated when enough restraints were found such that the 10 lowest energy structures converged to the average structure with a 1.5 Å RMSD. At this point, the structure was used to help resolve degeneracies in the spectra. Additionally, weak peaks in the proton NOESY spectra that did not have a corresponding heteronuclear chemical shift were included if the structure indicated a clear candidate proton. NOE restraints to aromatic protons were also included at this point. During the third phase, potential NOEs within a 4 Å radius were predicted based on the phase 2 structures and the data was examined again to further resolve degeneracies and incorporate more constraints.

NOE intensities were converted to distances by calibrating easily identifiable sequential NOE resonances that were clearly involved in a β-sheet. To account for spin-diffusion, the calculated distances were multiplied by a factor, which increased linearly from 1.0 to 1.3 as the interproton distance increased from 1.8 to 5 Å (*26*). Distance constraints were assigned error bars that increased linearly from ±15% at 1.8 Å to ±35% at 5 Å. These adjustments were made to account for the effects of spin-diffusion on the intensity of long-range distance constraints (*27*). Error bars for constraints to methyl groups and aromatic protons were increased by an additional 15% to account for the effects of rotation on the conversion of the NOE intensity to distance.

Also employed as constraints were three disulfide bonds and 35 hydrogen bonds. The disulfide bonds were previously determined in Der f 2 (*15*), and it was assumed that given the high sequence similarity the same disulfide bonds would be found in Der p 2. Amides, which were potentially involved in hydrogen bonds, were identified on the basis of their exchange kinetics. The amide of peaks that remained in the $^{1}H/^{15}N$-HSQC spectra 4 h after placing the protein in deuterated buffer at p^{2}H 6.0 were examined for a potential hydrogen bond acceptor. The hydrogen bond acceptor for these amides was identified by analysis of NMR structures at various stages of the refinement.

Main-chain (ϕ) dihedral constraints were determined from the J_{HNHA} coupling constants which were measured in the

[1] Abbreviations: NMR, nuclear magnetic resonance; ELISA, enzyme-linked immunosorbent assay; RAST, radio-allergosorbent test; RMSD, root mean squared deviation; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser enhancement; mAb, monoclonal antibody.

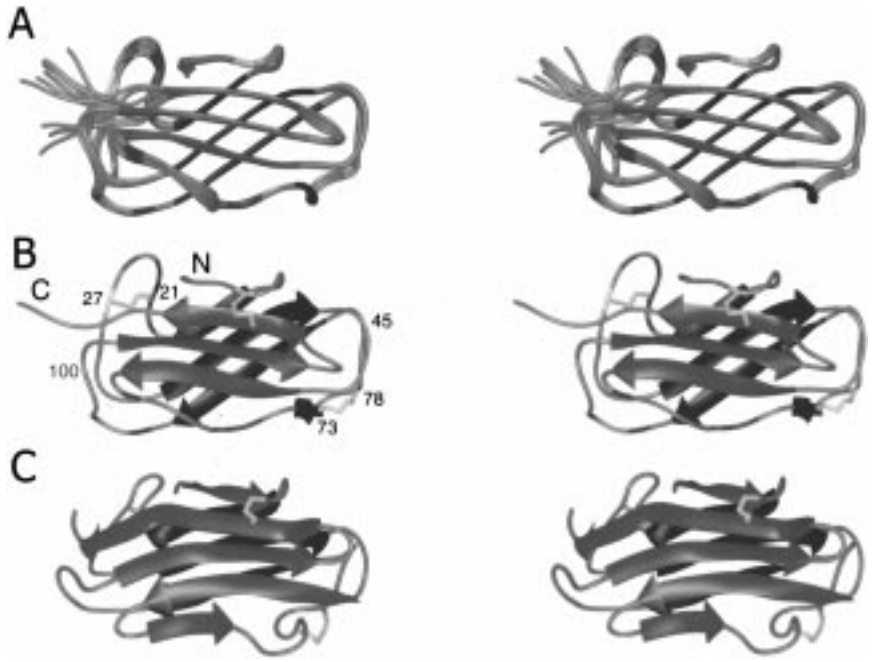

FIGURE 1: Stereo renderings of the NMR Structure of Der p 2 and Der f 2. Panel A shows the 10 lowest energy structures of Der p 2. Panel B shows a simplified ribbon diagram of the lowest energy structure of Der p 2; the annotations indicate the N- and C-terminus as well as the residue number of some cysteine residues as well as other residues (45 100) known to affect antibody binding to Der p 2. The disulfide bond 8−119 is shown unlabeled because the numbering can be interfered from the image. The color-coding in each structure is red and blue, secondary structural element; yellow, cysteine; green, residues determined to have multiple conformations. The structure is composed of two-three stranded antiparallel β-sheets and a short helix (2_7 ribbon) shown in blue. Regions that are less well constrained frequently contain residues which were previously determined to have multiple conformations; shown in green (*22*). Panel C shows the Der f 2 structure (PDB filename 1ahk) with the previously reported secondary structure elements (*21*). It is displayed aligned with Der p 2. These renderings were created using the program MIDAS (UCSF Computer Graphics Laboratory).

HNHA experiment (*28*). During the first phase of structure building, only residues with a single possible dihedral value based on the Karplus curve were incorporated. During the third phase of building, residues were assigned a ϕ value if the dihedral angle orientation in 10 of the 10 lowest energy structures was consistent with one of the possible dihedral angles. The acceptable range of dihedral constraints was initially plus or minus 30°, and subsequently tightened to 10° in the third phase if the majority of the structures showed a single value for ϕ. Side-chain dihedral angles were obtained from an analysis of the J_{NHB} coupling constant [obtained from the HNHB experiment (*29*)] and by examining H_N and H_α to H_β NOE intensities. The final structure included 88 ϕ constraints, 21 χ_1 constraints, and for Leu17, χ_2.

The structures were generated with X-PLOR version 3.1 using a distance geometry/simulated annealing protocol with multiple cycles of refinement (*30*). The constraints used in the distance geometry embedding protocol included backbone and aliphatic side-chain protons. Regularization of the embedded structures was done using the square well NOE potential and the refinement of these structures was done using the biharmonic potential. One hundred structures were embedded, regularized, and refined. The 10 lowest energy structures were further refined with an additional 20 cycles of simulated annealing where the first refinement cycle began at 2000K and final cycle at 1000K. Simultaneously, the weighting of the dihedral constraints increased from 100 to 200. The structures were analyzed by PROCHECK (*31*) and aligned using ALIGN (*32*). The dssp algorithm by Kabsch and Sander evaluated the structures for elements of secondary structure (*33*). The solvent accessible surface of the lowest energy structures of Der p 2 were evaluated with a 9.0 Å probe using the program MS (*34*).

Heteronuclear NOEs were measured using previously published sequences (*35*), and the ratio of the cross-peak intensity with and without presaturation is reported. A ratio of 0.78 indicates that the amide group shows no internal mobility. As this ratio decreases, the degree of internal mobility increases; disordered regions show a negative ratio.

Structure and Sequence Comparison. The 10 lowest energy structures were submitted for comparison to the DALI server (*36*, *37*). DALI provides the significance of the match given as a Z-score and a root-mean-squared deviation (RMSD) of the three-dimensional alignment. The Z-score is a composite of several evaluations including RMSD and number of residues matched. A score less than 2.0 is not a significant match, while a very strong match is typically greater than 10.0. Sequence homologies to the Der p 2 sequence were obtained using the program FASTA (*38*).

RESULTS

Structure Determination. The ensemble of the 10 lowest energy structures of Der p 2 is shown in Figure 1A. The structure contains two three-stranded antiparallel β-sheets (Figure 1B). Sheet 1, colored red, appears in the foreground and comprises residues 51−58, 104−111, and 118−122. In the background, sheet 2, which is colored blue, comprises residues 15−17, 35−44, and 84−92. In three dimensions, the two sheets overlay each other at an angle of approximately 30°, which is characteristic of the immunoglobulin fold. The two three-stranded antiparallel β-pleated sheets are slightly shorter than previously reported (*22*).

Table 1: Constraints and Summary of Structural Statistics[a]

number of long-range NOEs	414
number of short-range NOEs	956
number of hydrogen bonds	35
RMSD all backbone atoms (Å)[C]	0.93
RMSD of β-strands (Å)[c]	0.33
RMS bonds (Å)[b]	0.01
RMS angles (deg)[b]	1.55
RMS impropers[b]	1.26

NOE violations		dihedral violations	
distance (Å)	no. of violations	angle (deg)	no. of violations
0.3	13.8	3.0	8.3
0.4	2.9	5.0	0.3
0.5	0.0	7.0	0.0

	ϕ/ψ space, percent residues[d]
most favored region	58.0
additionally allowed region	35.1
generously allowed region	5.6
disallowed region	1.3

[a] Calculations were done on the 10 lowest energy structures displayed in Figure 1A. The numbers of NOE and Dihedral violations are averaged over the 10 structures for the distance or angle cutoff displayed. Short-range NOEs are defined as intraresidue or sequential, while all others are long range. [b] Calculated with X-PLOR (*30*). [c] Calculated with ALIGN (*32*). [d] Calculated with PROCHECK (*31*).

Preliminary analysis predicted β-strands for residues 45−48 and 80−82 but these residues are simply two loops in close association and are not identified as β-strands by dssp. There is also a short helix (2_7 ribbon) that runs from residues 72 to 75.

These structures show a 0.93 Å RMSD alignment for all backbone atoms while the backbone atoms of secondary structural elements align to each other with a RMSD of 0.33 Å. The structures show good local geometry and have an equivalent resolution to X-ray derived structures of 3.3 Å as determined with PROCHECK. A summary of the structural statistics can be found in Table 1. The NOE and dihedral violations are evaluated at different distance and angle thresholds in order to give an impression of the distribution of errors. There are no NOE errors greater than 0.5 Å or dihedral errors greater than 7°. An examination of the number of the predicted NOEs within a 4.5 Å radius indicated that approximately one-third of the available NOEs had been incorporated as constraints (data not shown).

Figure 2 displays statistics related to the quality of the structure on a per residue basis. The secondary structure is indicated at the top of the figure with filled in boxes, and the open boxes show residues that were previously described to have multiple conformations. The figure displays the RMSD to the average structure (panel A), number of long range NOEs (panel B), and heteronuclear NOE ratio (panel C). Regions of the protein that show the lowest RMSD participate in the formation of secondary structure. Regions that show a high RMSD (i.e. the regions between $\beta1-\beta2$, $\beta3-\alpha$, and $\beta6$-carboxy terminus) may do so either because of a lack of constraints in these regions or due to the presence of internal motion. The data presented in Figure 2B indicate that these regions have fewer long-range NOEs. The ratios of the peak intensities from the heteronuclear NOE experiments (Figure 2C) are relatively homogeneous and a discernible trend is not obvious. This indicates that the protein is likely not undergoing rapid time scale motions. Conseqently, the higher RMSD for these regions is likely to be a consequence of the number of constraints. The reduced number of constraints may be due to low-frequency motions. Some of the regions with higher RMSDs are nearby to residues that were previously described to exist in multiple conformations. This is most evident in the residues surrounding Val 94 and Ile 97. Therefore, it is possible that these regions of the protein may be undergoing motion on a slower time scale. Such motions may attenuate the intensity of NOE peaks due to exchange broadening as well as the existence of multiple peaks for some resonances.

Homologies. In terms of sequential homology, Der p 2 matches with the similar group 2 proteins from other mite species, Der f 2 (87% identity) and Lep d 2 (40% identity). The alignment of Der p 2 with Der f 2 is shown in Figure 3. Der p 2 was previously shown have 29% sequence identity with HE1, a human epididymal gene product suggesting that the group 2 allergens may play a role in mite reproduction (*7*). There is a 25% identity between Der p 2 and for the chimp homologue of HE1 (CE1). We also found in a FASTA (*38*) search that Der p 2 shares 25% sequence identity and 55% homology with the esr16 protein of moths (*39*, *40*). The Z-score is 198.8 and Smith−Waterman score (*41*) is 137, which indicates a significant match.

To explore the structural homology of Der p 2 to other possible folds, the program DALI examined the 10 lowest energy structures (*36*, *37*). The results are summarized in Table 2 where the significance (Z-score) and RMSD are reported for those structures which matched all 10 of the lowest energy NMR structures. As expected, Der p 2 matches rather strongly with Der f 2, Z-score 8.6 and RMSD 4.9. However, there is another set of structures with lower Z-scores but slightly better alignments. Fourteen other structures matched with average Z-scores between 2.4 and 4.4 and the average alignments ranged from 2.9 to 4.2 Å RMSD. Of the low Z-score structures, seven are found in the Structural Classification of Proteins (SCOP) database (*42*). These seven fall into two different superfamilies of the immunoglobulin-like β-sandwich fold: immunoglobulin and fibronectin type III.

Structral Differences between Der p2 and Der f2. Der p 2 and Der f 2 share a high-sequence homology (see Figure 3A) as well as a high similarity in secondary structure but curiously do not show a very high structural homology. Figure 1 juxtaposes the aligned structures of Der p 2 (panel B) and Der f 2 (panel C). Some structural features are similar such as the positioning and orientation of the termini and the short loops pinched off by the disulfide bonds. The major differences in the two structures are the reported number of β-strands and the orientation of the two sheets with respect to each other. The differences in the β-strands appear to be due to the method used to determine secondary structure. For example, in Der f 2, residues 5−8 and 62−65 are reported to be β-sheet, whereas they are not in Der p 2. The secondary structure of these β-strands in Der f 2 was determined from NOEs and chemical shifts. A similar method when applied to Der p 2 also shows a secondary structure that is similar to Der f 2 in these regions. However, a more rigous and unbiased analysis using dssp does not assign residues 5−8 or 62−65 to β-strand in either Der f 2 or Der p 2.

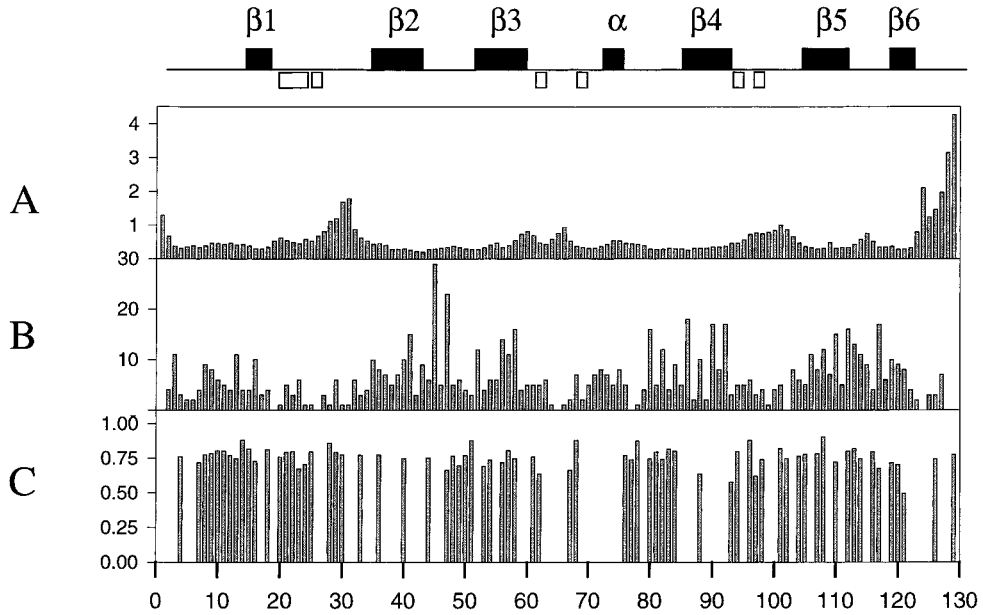

FIGURE 2: Summary of Structural Statistics Per Residue. The top of the figure shows the location of secondary structure (closed boxes) and the location of residues known to undergo conformation exchange (open boxes). (A) The RMSD of the backbone atoms with respect to the average structure are shown per residue. (B) The number of long-range NOEs is shown for each residue with long-range being defined as between residues that are more than 1 away in primary sequence. (C) The heteronuclear (proton-nitrogen) NOE is shown. The absense of a bar indicates that the amide resonance from that residue was not sufficiently resolved to measure the intensity in the spectra. An immobile residue would show a maximum of 0.78 for the heteronuclear NOE. A value of less than zero indicates extensive motion in the nanosecond time scale. None of the residues in Der p 2 showed a negative NOE.

The difference in the orientation of the two β-strands in Der p 2 and Der f 2 is significant; in Der p 2, the two sheets are oriented at approximately 30°, while in Der f 2, the strands of the sheets are closer to parallel. The interstrand NOEs that define the sheets for Der p 2 and Der f 2 are virtually identical, indicating that the residues that interact to form the individual strands are virtually identical in both proteins (*21*, *22*). For example, in both the Der p 2 and Der f 2 structures, residues 38−43 run antiparallel with residues 83−92. This represents a significant agreement between the two structures so it does not seem likely that the strand alignment is the cause of the differences in the two structures.

To investigate whether the solution conditions were likely responsible for the differences between the structures of Der f 2 and Der p 2, we compared the HSQC spectra of Der p 2 in the different solution conditions used for structure determination. Figure 4 shows the HSQC for Der p 2 at 25 C in P2 buffer (Figure 4A), at 55 °C in P2 buffer (Figure 4B) and finally at 55 °C in F2 buffer (Figure 4C). When comparing panel A to panel B, it appears that some amides shift slightly in response to temperature. In contrast, there are a large number of shifts when the protein was placed in the F2 buffer, which contains *N*-octyl-β-D-glucoside (compare panels B and C of Figure 4). Although some of the cross-peaks in the less crowded areas of the spectrum do not shift significantly, in the central region of the spectrum there are numerous shifts that render the previous assignments impossible to interpret. This indicates that many amides have significantly changed their chemical environment due to the presence of the detergent.

DISCUSSION

Comparisons of Der p 2 to Der f 2. Der p 2 and Der f 2 are highly similar in terms of sequential identity yet the structures appear very different. To examine the possibility that the sequence could account for the structural differences, the residues that are not identical in sequential alignment were examined for their position in the three-dimensional structure (see Figure 3A). Eleven of the 14 substitutions occur in loops in Der p 2. Most of the differences are conservative substitutions as evaluated by FASTA (*38*). It is possible that the subtle differences in residue type in the loops change the orientation of the sheets but this is regarded as unlikely.

What is most likely the source of the structural differences shown here is the different conditions in which the structures were determined. The buffer used here for Der p 2 is composed of various salts at a total concentration that is physiological. In addition, the temperature was maintained at 25 °C for the NMR experiments. This is contrasted with the buffer used for Der f 2 which contained no salt and the detergent *N*-octyl-β-D-glucoside at 140 mM. Furthermore, the Der f 2 NMR experiments were carried out at 55 °C. The critical micelle concentration for *N*-octyl-β-D-glucoside is 20 mM (*43*), so it is likely that Der f 2 was in the micelles. Our examination of the HSQC spectrum of Der p 2 under these various solution conditions (Figure 4) indicated that many amides experienced a different chemical environment when in the different buffers. We propose that these shifts strongly suggest that the differences in the structure are solely due to the solution conditions used in the structure determination and that the Der p 2 structure was determined under more physiologically relevant conditions.

The Der p 2 Structure. The results presented here indicate that the topology of Der p 2 is well determined by NMR methods. Overall, the structure is cylindrical with one end of the cylinder exhibiting more conformational diversity than

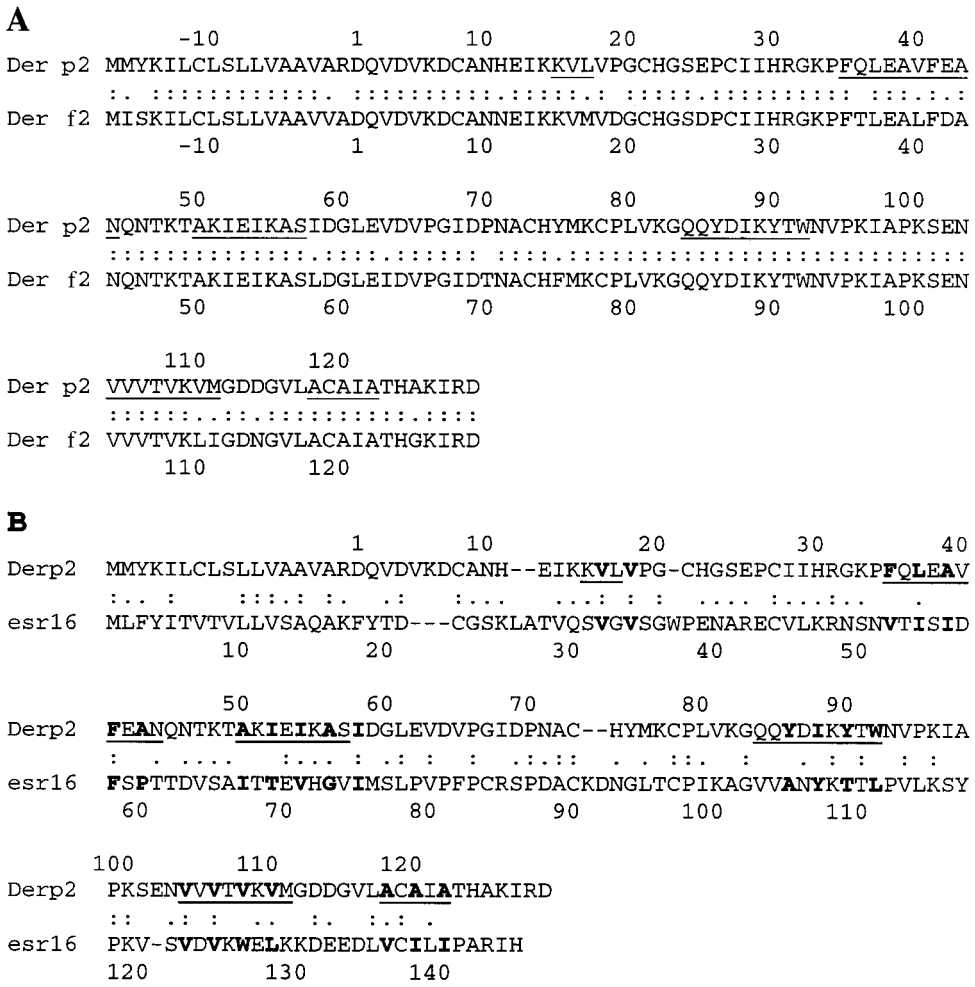

FIGURE 3: Sequential Alignments of Der p 2. The sequential alignment of the protein Der p 2 (including a leader sequence of 18 residues) with Der f 2 (Panel A) and esr16 (Panel B) are shown as determined by FASTA (*38*). A colon (:) indicates identity while a period (.) indicates homology. In the case of Der p 2 and Der f 2 the residues are numbered such that the amino-terminal residue in the mature protein is residue 1. Residues in the secondary structural elements of Der p 2 are shown underlined. In the comparison of Der p 2 with esr16 the buried residues which are in or adjacent to β-strands are indicated in bold type. Der f 2: Z-score, 1302.8; Smith−Waterman score, 881; 86.99% identity in 146 aa overlap. esr16: Z-score, 198.8; Smith−Waterman score, 137; 25.87% identity in 143 amino acid overlap.

the other (Figure 1A). The poorly constrained regions at one end of the cylinder contain residues previously suggested to have multiple conformations based on multiple peaks in the HSQC spectrum (*22*). These residues are highlighted in green in Figure 1A. In other proteins, slow conformational changes are often associated with a cis−trans isomerization of proline residues (*44*, *45*). An examination of sequential NOE data for the proline residues in Der p 2 indicated that eight of the nine prolines were in the trans conformation and Pro 79 was in a cis conformation. However, it was not apparent from the NOE data that there was any proline cis−trans isomerization. Therefore, we cannot explain the origin of the multiple conformations at this time, but they are likely the cause of the conformational diversity at the one end of the molecule.

Structural Homology to the Immunoglobulin Fold. The observed orientation of the two sheets in Der p 2 is characteristic of the immunoglobulin fold and the sequential order in which the strands interact is similar to that of this family of proteins (*22*). However, Der p 2 has only six strands compared to seven in most proteins with the immunoglobulin folds. The crucial strand that would define the "Greek-key" should be located between amino acids 58−72. Visual inspection of the structural alignments reveals that the residues comprising the "missing strand" in Der p 2 align with the predicted strand from the immunoglobulin fold. The reason these residues are not found in a β-strand in Der p 2 is that this region contains two proline and two glycine residues. These residues probably disrupt secondary structure and prevent this strand from forming the correct hydrogen bonding. Der p 2 shows low Z-scores and poor RMSD alignment with members of the immunoglobulin family (Table 2) in two major folds: immunoglobulin and fibronectin type III. This may suggest that Der p 2 represents a new submember of the immunoglobulin-like β-sandwich family. However, the relatively low resolution of the NMR structure of Der p 2 (3.3 Å) may also account for the low Z-scores.

Implications of the Sequential Homology to esr16. Since the three-dimensional structure search did not find a distinct homologue of Der p 2 with enzymatic activity, the sequential homologies were reexamined. Der p 2 was found to match a recent entry in the database, esr16. In addition to the sequential alignment, it is likely that Der p 2 and esr16 are similar in tertiary structure because most of the residues that form the hydrophobic core of Der p 2 are conserved in esr16 (see Figure 3B).

Table 2: Output of DALI Examination of Possible Structural Homologies and SCOP Classification[a]

PDB name	avg Z-score	avg RMSD	SCOP classification	source protein
1ahk	8.58	4.85		Der f 2
1fie (A)	4.43	3.36	Igβ:Ig:E-set	coagulation factor XIII
1eut	3.66	3.38	Igβ:Ig:E-set	Sialidase
1axi (B)	3.54	2.91		growth hormone receptor
1cdy	2.97	3.40		Cd4
8fab (A)	3.33	3.38	Igβ:Ig:V-set	IgG1
1ten	3.01	3.09	Igβ:F3:F3	Tenascin
1ebp (A)	3.34	3.27		epo receptor
1fnf	3.00	3.34	Igβ: F3: F3	Fibronectin
1iak (B)	2.95	3.46		mhc-class II
1zxq	2.63	3.60		intercellular adhesion molecule 2
1cfb	2.64	3.39	Igβ:F3:F3	Drosophila neuroglian
1igt (B)	2.61	3.80		IgG2a
1fc1 (A)	2.49	3.57	Igβ:Ig:C1-set	Fc of IgG1
1cd1 (A)	2.44	4.22		cd1

[a] DALI examined the 10 lowest energy structures for possible structural homologies. The average Z-score, average RMSD, and name of the protein are shown. The SCOP database was searched for proteins that have been classified according to their fold. The classification, if found, is listed as fold:superfamily:family. The abbreviations are Igβ, immunoglobulin-like β-sandwich; Ig, immunoglobulin; F3, fibronectin type III.

The homology between Der p 2 and esr16 is striking in terms of a possible function of the Der p 2 protein in mites. The esr16 protein was identified in the moth *Manduca sexta* and localizes to the tracheal epithelial cells (*40*). Prior to pupal and larval ecdysis, esr16 mRNA levels increase and the expression can be regulated with ecdysteroids (*39*). This suggests that esr16, and possibly Der p 2, is involved in molting. The fecal particles of dust mite are similar in size to grass pollen particles and are small enough to become air-borne under standard household conditions but, whole dust mites are not (*46*). It is not well understood why Der p 2 occurs mostly in extracts of the mite whole body and yet is a major inhalant allergen (*4*). We speculate that Der p 2 could be found in the shed exoskeleton of mites (during molting) and can disintegrate into smaller particles that become air-borne.

Implications of the Structure to Allergy and Immunotherapy. The structure presented here explains the involvement of residues 44−46 and 100 in Der p 2 and the role of various residues in Der f 2 in antibody binding. In all cases, these residues are found on the surface of the protein (see Figure 1B) and thus would be available for interaction with antibody. The Der p2 structure also explains previous efforts to engineer group 2 allergens to reduce the risk of systemic anaphylaxis by eliminating one or more disulfide bonds (*18*−*20*). The three disulfide bonds in Der p2 and Der f2 have been shown to be important for antibody binding. The disulfide bonds 21−27 and 73−78 are found in loops at almost opposite ends of the molecule. The 8−119 bond is more central to the body of the structure and links the amino-terminus to one of the two β-pleated sheets (Figure 1B). Although these disulfide bonds are near the surface of the molecule, they have little to no solvent-accessible surface evaluated by examining all 10 structures with a 9.0 Å probe using the program MS (*34*). Therefore, the reduced IgE-mediated skin reactivity of the Cys mutants in the previous studies cannot be attributed to a direct role of the cysteine side-chain interaction with IgE. Rather, disruption of one or more of these disulfide bonds would lead to local or global conformational alterations in Der p 2 and subsequent loss of recognition by IgE.

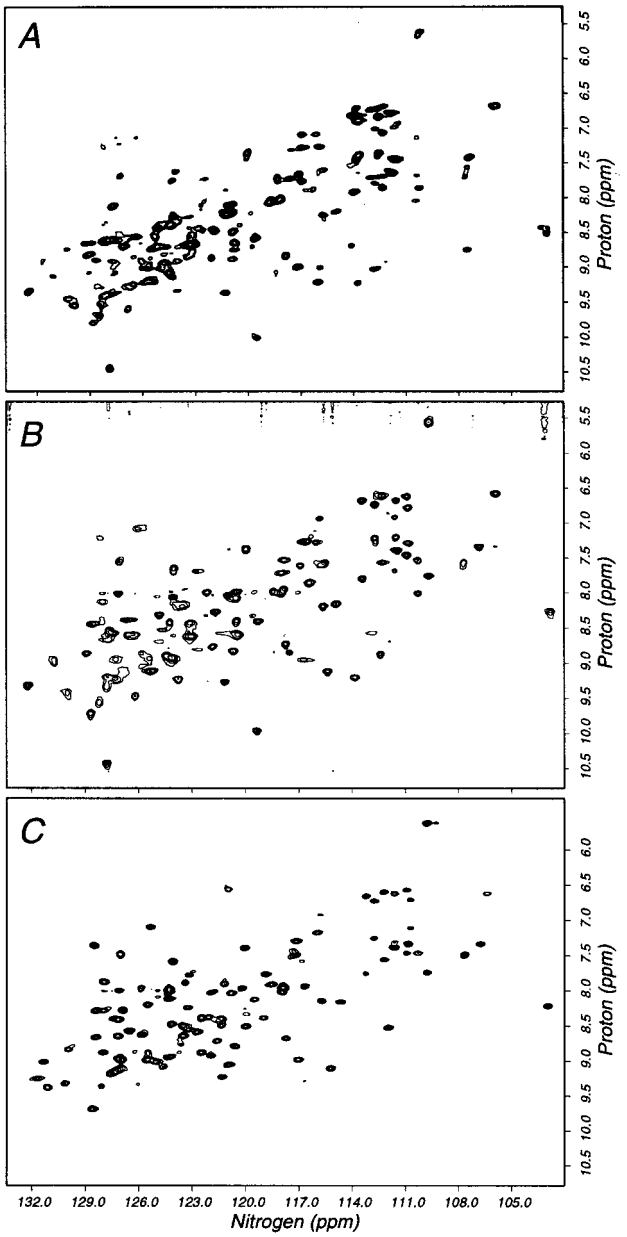

FIGURE 4: Der p 2 in different solution conditions. An $^1H/^{15}N$−HSQC of Der p 2 was taken under various solution conditions: (A) p2 buffer, 25 C; (B) p2 buffer, 55 C; (C) f2 buffer, 55 C. P2 buffer: 10 mM sodium phosphate, pH 6.0, 50 mM NaCl, 100 mM K_2SO_4, and 5 μM EDTA. F2 buffer: 140 mM *N*-octyl-β-D-glucoside, pH 5.6.

These engineered allergen studies clearly show that patient IgE-mediated skin test reactivity is significantly reduced using such recombinant allergens while T-cell proliferative responses in vitro remain the same as for natural Der p 2 (*18*, *19*). Various studies indicate that patients see a vast array of T-cell epitopes and that no one epitope that is "immunodominant" for Der p 2 (reviewed in ref *11*). Therefore, modified allergen is an ideal immunotherapeutic agent because it would present a broad spectrum of possible T-cell epitopes to patients with various MHC alleles yet reduce IgE mediated responses. While the reduction of

disulfide bonds in allergens is clearly of value in activating regulatory T-cells while preventing IgE-mediated responses, this strategy is limited to those allergens possessing such disulfide bonds. We suggest that, given the three-dimensional structure of an allergen, an alternative and more general approach to allergen immunotherapy would be to alter surface residues to reduce or eliminate deleterious IgE-mediated reactivity during desensitization without affecting the activation of regulatory T-cells.

ACKNOWLEDGMENT

The authors would like to thank A. M. Smith, M. D. Chapman, and T.A.E. Platts-Mills for their expertise and encouragement. We thank T. C. Wood for performing the sequence analysis and helping in the interpretation of the homology data.

REFERENCES

1. Platts-Mills, T. A., Vervloet, D., Thomas, W. R., Aalberse, R. C., and Chapman, M. D. (1997) *J. Allergy Clin. Immunol. 100*, S2−24.
2. Heymann, P. W., Chapman, M. D., Aalberse, R. C., Fox, J. W., and Platts-Mills, T. A. (1989) *J. Allergy Clin. Immunol.* 83, 1055−1067.
3. Sporik R., Chapman M. D., and Platts-Mills T. A. (1992) *Clin. Exp. Allergy 22*, 897−906.
4. Platts-Mills, T. A., Thomas, W. R., Aalberse, R. C., Vervloet, D., and Champman, M. D. (1992) *J. Allergy Clin. Immunol. 89*, 1046−1060.
5. Thomas, W. R. (1996) in *Molecular Analysis of Immediate Hypersensitivity and Strategies for Immunological Intervention* (Walker, M. R., and Roberts, A. M., Eds.) John Wiley & Sons, Chichester, U.K.
6. Robinson, C., Kalsheker, N. A., Srinivasan, N., King, C. M., Garrod, D. R., Thompson, P. J., and Stewart, G. A. (1997) *Clin. Exp. Allergy 27*, 10−21.
7. Thomas, W. R., and Chua, K. Y. (1995) *Clin. Exp. Allergy 25*, 667−669.
8. Ferreira, F., Ebner, C., Kramer, B., Casari, G., Briza, P., Kungl, A. J., Grimm, R., Jahn-Schmid, B., Breiteneder, H., Kraft, D., Breitenbach, M., Rheinberger, H. J., and Scheiner, O. (1998) *FASEB J. 12*, 231−242.
9. Burks, A. W., Shin, D., Cockrell, G., Stanley, J. S., Helm, R. M., and Bannon, G. A. (1997) *Eur. J. Biochem. 245*, 334−339.
10. Stanley, J. S., King, N., Burks, A. W., Huang, S. K., Sampson, H., Cockrell, G., Helm, R. M., West, C. M., and Bannon, G. A. (1997) *Arch. Biochem. Biophys. 342*, 244−253.
11. Norman, P. S. (1993) *Curr. Opin. Immunol. 5*, 968−973.
12. van't Hof, W., Driedijk, P. C., van, d. M., Beck-Sickinger, A. G., Jung, G., and Aalberse, R. C. (1991) *Mol. Immunol. 28*, 1225−1232.
13. Chua, K. Y., Greene, W. K., Kehal, P., and Thomas, W. R. (1991) *Clin. Exp. Allergy 21*, 161−166.
14. Benjamin, D. C., Berzofsky, J. A., East, I. J., Gurd, F. R., Hannum, C., Leach, S. J., Margoliash, E., Michael, J. G., Miller, A., and Prager, E. M., et al. (1984) *Ann. Rev. Immunol. 2*, 67−101.
15. Nishiyama, C., Yuuki, T., Takai, T., Okumura, Y., and Okudaira, H. (1993) *Int. Arch. Allergy Immunol. 101*, 159−166.
16. Lombardero, M., Heymann, P. W., Platts-Mills, T. A., Fox, J. W., & Chapman, M. D. (1990) *J. Immunol. 144*, 1353−1360.
17. Nishiyama, C., Fukada, M., Usui, Y., Iwamoto, N., Yuuki, T., Okumura, Y., and Okudaira, H. (1995) *Mol. Immunol. 32*, 1021−1029.
18. Takai, T., Yokota, T., Yasue, M., Nishiyama, C., Yuuki, T., Mori, A., Okudaira, H., and Okumura, Y. (1997) *Nat. Biotechnol. 15*, 754−758.
19. Smith, A. M., and Chapman, M. D. (1996) *Mol. Immunol. 33*, 399−405.
20. Smith, A. M., and Chapman, M. D. (1997) *Clin. Exp. Allergy 27*, 593−599.
21. Ichikawa, S., Hatanaka, H., Yuuki, T., Iwamoto, N., Kojima, S., Nishiyama, C., Ogura, K., Okumura, Y., and Inagaki, F. (1998) *J. Biol. Chem. 273*, 356−360.
22. Mueller, G. A., Smith, A. M., Williams, D. J., Hakkaart, G. A., Aalberse, R. C., Chapman, M. D., Rule, G. S., and Benjamin, D. C. (1997) *J. Biol. Chem. 272*, 26893−26898.
23. Bodenhausen, G., and Ruben, D. G. (1993) *Chem. Phys. Lett. 69*, 185−189.
24. Brown, S. C., Weber, P. L., and Mueller, L. (1988) *J. Magn. Reson. 71*, 166−171.
25. Briercheck, D. M., Wood, T. C., Allison, T. J., Richardson, J. P., and Rule, G. S. (1998) *Nat. Struct. Biol. 5*, 393−399.
26. Suri, A. K., and Levy, R. M. (1993) *J. Magn. Reson., Ser. B 101*, 320−324.
27. Borgias, B. A., Gochin, M., Kerwood, D. J., and James, T. L. (1990) *Prog. NMR Spectrosc. 22*, 83−100.
28. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
29. Archer, S. J., Ikura, M., Torchia, D. A., and Bax, A. (1991) *J. Magn. Reson. 95*, 636−641.
30. Brunger, A. (1993) in *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven,.
31. Laskowski, R. A., Rullman, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1998) *J. Biomol. NMR 8*, 477−486.
32. Satow, Y., Cohen, G. H., Padlin, E. A., and Davies, D. R. (1996) *J. Mol. Biol. 190*, 593−604.
33. Kabsch, W., and Sander, C. (1983) *Biopolymers 22*, 2577−2637.
34. Connoly, M. L. (1983) *Science 221*, 709−713.
35. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.
36. Holm, L., and Sander, C. (1993) *J. Mol. Biol. 233*, 123−138.
37. Holm, L., and Sander, C. (1998) *Nucleic Acids Res. 26*, 316−319.
38. Pearson, W. R. (1991) *Genomics 11*, 635−650.
39. Meszaros, M., and Morton, D. B. (1996) *Insect Biochem. Mol. Biol. 26*, 7−11.
40. Meszaros, M., and Morton, D. B. (1996) *J. Exp. Biol. 199*, 1555−1561.
41. Smith, T. F., and Waterman, M. S. (1981) *J. Mol. Biol. 147*, 195−197.
42. Murzin, A. G., Brenner, S. E., Hubbard, T., and Clothia, C. (1995) *J. Mol. Biol. 247*, 536−540.
43. Brito, R. M., and Vaz, W. L. C. (1986) *Anal. Biochem. 152*, 250.
44. Evans, P. A., Dobson, C. M., Kautz, R. A., Hatfull, G., and Fox, R. O. (1987) *Nature 329*, 266−268.
45. Chazin, W. J., Kordel, J., Drakeberg, T., Thulin, E., Brodin, P., Grundstrom, T., and Forsen, S. (1989) *Proc. Natl. Acad. Sci. 86*, 2195−2198.
46. Tovey, E. R., Chapman, M. D., and Platts-Mills, T. A. (1981) *Nature 289*, 592−593.

BI980578+