Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 29, Issue Number 4, (2012)

Modelling and Bioinformatics Analysis of the Dimeric Structure of House Dust Mite Allergens from Families 5 and 21: Der f 5 Could Dimerize as Der p 5

http://www.jbsdonline.com

Souad Khemili[1]*
Jean Marc Kwasigroch[2]
Tarik Hamadouche[1]
Dimitri Gilis[2]

[1]Département de Biologie, Université M'Hamed Bougara de Boumerdes, Algeria.
[2]Unité de Bioinformatique génomique et structurale, Université Libre de Bruxelles, Brussels, Belgium.

Abstract

Allergy represents an increasing thread to public health in both developed and emerging countries and the dust mites *Dermatophagoides pteronyssinus* (Der p), *Blomia tropicalis* (Blo t), *Dermatophagoides farinae* (Der f), *Lepidoglyphus destructor* (Lep d) and *Suidasia medanensis* (Sui m) strongly contribute to this problem. Their allergens are classified in several families among which families 5 and 21 which are the subject of this work. Indeed, their biological function as well as the mechanism or epitopes by which they are contributing to the allergic response remain unknown and their tridimensional structures have not been resolved experimentally except for Blo t 5 and Der p 5. Blo t 5 is a monomeric three helical bundle, whereas Der p 5 shows a three helical bundle with a kinked N-terminal helix that assembles in an entangled dimeric structure with a large hydrophobic cavity. This cavity could be involved in the binding of hydrophobic ligands, which in turn could be responsible for the shift of the immune response from tolerance to allergic inflammation. We used molecular modelling approaches to bring out if other house dust mite allergens of families 5 and 21 (Der f 5, Sui m 5, Lep d 5, Der p 21 and Der f 21) could dimerize and form a large cavity in the same way as Der p 5. Monomeric models were first performed with MODELLER using the experimental structures of Der p 5 and Blo t 5 as templates. The ClusPro server processed the selected monomers in order to assess their capacity to form dimeric structures with a positive result for Der p 5 and Der f 5 only. The other allergens (Blo t 5, Sui m 5, Lep d 5, Der p 21 and Der f 21) did not present such a propensity. Moreover, we identified mutations that should destabilize and/or prevent the formation of the Der p 5 dimeric structure. The production of these mutated proteins could help us to understand the role of the dimerization process in the allergic response induced by Der p 5, and if Der p 5 and Der f 5 behave similarly.

Key words: House dust mite allergens; Families 5 and 21; Comparative modeling; Protein-protein docking.

Introduction

House dust mites are an important source of allergens that cause asthma, eczema and rhinitis. Mite species *Dermatophagoides pteronyssinus* (Der p) and *Dermatophagoides farinae* (Der f) are found worldwide, whereas the *Blomia tropicalis* specie (Blo t) is mainly found in tropical regions. Allergens from dust mites are quite diverse, with more than 20 groups identified so far. The tridimensional (3D) structures of some of these allergens, namely Der p 1 (1, 2), Der p 2 (3), Der p 5 (4), Der p 7 (5), Der f 1 (6), Der f 2 (7), Der f 13 (8), Blo t 5 (9, 10), have been solved by X-ray crystallography or by nuclear magnetic resonance (NMR). These 3D structures should provide a better understanding of the biological function of the corresponding proteins and of their capacity to bind some hydrophobic ligand that could provoke Th2-type immune response (11). Together with the identification of the epitopes

*Corresponding author:
Souad Khemili
Phone: +213 771 426 506
E-mail: skhemili@ulb.ac.be
jkwasigr@ulb.ac.be
tarikhamadouche@gmail.com
dgilis@ulb.ac.be

that are responsible for the allergic response, 3D structures should help the design of hypoallergens for specific immunotherapy.

Group 5 allergens have been isolated from several species: Der p 5 from *Dermatophagoides pteronyssinus*, Blo t 5 from *Blomia tropicalis*, Der f 5 from *Dermatophagoides farinae*, Lep d 5 from *Lepidoglyphus destructor*, Ale o 5 from *Aleuroglyphus ovatus*, Tyr p 5 from *Tyrophagus putrescentitiae*, Sui m 5 from *Suidasia medanensis* and Gly d 5 from *Glycyphagus domesticus*. The sequences of these allergens are highly analogous to those of the group 21 (Der p 21, Der f 21 and Blo t 21). However, few information regarding their biological function, their potential cross-reactivity and their epitopes is available.

The 3D structure of the allergens from groups 5 and 21 have not yet been determined experimentally except for Blo t 5 (9, 10) and Der p 5 (4). The solution structure of Blo t 5 (PDB code 2JMH) shows a flexible 17 residue amino terminus followed by three similarly sized α-helices, which are packed into an antiparallel bundle (9). The hydrophobic side chains are tightly packed and the exchange of these interactions confers flexibility to the bundle. The X-ray structure of Der p 5 has recently been resolved (4) (PDB code 3MQ1) and is a three helical bundle similar to Blo t 5. Contrasting with Blo t 5 that is reported to be a monomer, the crystallographic asymmetric unit of Der p 5 contains three dimers. Der p 5 presents a concentration-dependent oligomerization. The dimerization process creates a large hydrophobic cavity that could be a ligand-binding pocket and that could be involved in the allergic response. In the dimeric structure of Der p 5, the N-terminal helix is kinked around the glycine at position 45 of the PDB (3MQ1); this glycine appears to be important for the flexibility of the N-terminal helix. The kink in the helix opens the structure of each monomer and leads to a dimer where both chains are entangled. This glycine is also present in Blo t 5, but the molecule is monomeric in the crystal and possesses a straight N-terminal helix. According to Mueller *et al.* (4), the sequence differences between Blo t 5 and Der p 5 at PDB positions 85 and 88 (PDB numbering of 3MQ1) might explain the fact that Der p 5 is mainly a dimer and Blo t 5 a monomer. Residue Val 88 of one chain of Der p 5 interacts with Val 85 of another chain, making a "valine zipper". Both positions are occupied by an Ala in Blo t 5 what should decrease the hydrophobic surface between the two chains and could explain that Blo t 5 remains monomeric.

Der p 21 shares 31% sequence identity with Der p 5. Its 3D structure has not been resolved, but has been studied by small angle X-ray scattering and circular dichroism (12). The results of these experiments suggest an α-helical secondary structure and a dimeric structure with an elongated shape. *Ab initio* simulations propose a dimeric structure that is different from Der p 5.

What is the dimerization propensity of allergens from families 5 and 21, what is the structure of the possible dimers, does the dimerization process create an hydrophobic cavity similar to that observed in Der p 5, ... are the questions we adressed. Indeed, the hydrophobic cavity created in a Der p 5 dimeric-like structure could accommodate some hydrophobic ligands that could be responsible for the shift of the immune response from tolerance to allergic inflammation. We tackled this problem by bioinformatics approaches and molecular modelling techniques. We focused this study on Blo t 5, Der p 5, Der f 5, Sui m 5, Lep d 5, Der p 21 and Der f 21. The proteins chosen in group 5 present a Gly at position 45, except Lep d 5. Der p 21 and Der f 21 have a sequence identity larger than 30% with some members of group 5 and no glycine at position 45. For each allergen, we generated a first set of monomeric models using the 3D structure of Blo t 5 as template. These models present a three-helix bundle structure, with a straight N-terminal helix. We computed a second set of monomeric models based on the structure of one monomer of Der p 5. The N-terminal helix of these models is kinked. The latter models were then used to perform protein-protein docking simulations and to predict the potential dimeric structure of these allergens.

The structures of the dimers were analyzed and the potential mode of their formation is discussed. We would like to stress that the softwares used to predict the possible dimeric structures of these proteins do not indicate whether they would actually form or not but instead whether their structure would be similar to that of Der p 5 in that case.

Methods

Sequence Alignment

The sequences of the various allergens were obtained from the UniProtKB/SwissProt protein knowledgebase (http://www.expasy.ch/sprot/). The entry codes used are: O96870 for Blo t 5, P14004 for Der p 5, A8B8I1 for Der f 5, B2GM87 for Sui m 5, Q9U5P2 for Lep d 5, Q2L7C5 for Der p 21 and A8B8G7 for Der f 21. The multiple sequence alignments were performed with the ClustalW program (http://www.ebi.ac.uk/Tools/clustalw2/index.html) (13), with the default parameters. The pairwise alignments were generated with Blast2seq (14), using the Blosum62 scoring matrix. The sequence similarity corresponds to the percentage of amino acids for which the alignment score is positive.

Comparative Modelling

MODELLER 9v4 (15) was used to build the 3D models of the allergens by comparative modelling. The 3D structures of the templates were extracted from the PDB database (16). A first set of models was computed for Der f 5, Sui m 5, Lep d 5, Der p 21 and Der f 21, using the structure of Blo t 5 (PDB code 2JMH) as template. This template presents a three-helix bundle conformation. A second set of models was computed for Blo t 5, Der f 5, Sui m 5, Lep d 5, Der p 21 and Der f 21, using the chain A of the structure of Der p 5 (PDB code 3MQ1) as template. These models present a kink in the N-terminal helix.

In this study, 30 runs for each protein and each template were carried out using MODELLER standard parameters. The quality of our models was assessed by two types of energy functions: a semi-empirical force field (Gromos) and a knowledge-based energy function (Anolea). Steepest descent energy minimizations of the 30 models of each allergen were performed using Gromacs3.3 (17) with the GROMOS96 45a3 force field. The Anolea score (18) was also computed for each model after energy minimization. Anolea is a knowledge-based energy function that evaluates the non-local environment (NLE) of each heavy atom in the model. The NLE is defined as the set of all heavy atoms within the distance of 7 Å that belong to amino acids farther than 11 residues along the sequence in the analyzed polypeptide.

The Gromos and Anolea scores were transformed into z-scores (equation 1), for normalization purpose.

$$z-\text{score}_i = \frac{S_i - \langle S \rangle}{\sigma} \qquad [1]$$

where S_i is the score of the i-th model computed with Anolea or Gromos, $<S>$ is the average of the scores of the 30 models, and σ is the standard deviation of the distribution of the scores. For each model i, the average between the z-scores computed with Anolea and Gromos, $<z\text{-score}_i>$, was computed. We selected the models presenting an average z-score lower than -0.5 for further protein-protein docking simulations. The number of models that have been obtained with the chain A of Der p 5 as template (3MQ1_A) and that have been carried is: 8 for Blo t 5 (blot5_monomer$_{1\to8}$[3MQ1_A]), 5 for Der f 5 (derf5_monomer$_{1\to5}$[3MQ1_A]), 7 for Sui m 5 (suim5_monomer$_{1\to7}$[3MQ1_A]), 6 for Lep d 5 (lepd5_monomer$_{1\to6}$[3MQ1_A]), 7 for Der f 21 (derf21_monomer$_{1\to7}$[3MQ1_A]) and 8 for Der p 21

(derp21_monomer$_{1\rightarrow8}$[3MQ1_A]). The number of models that have been obtained with the Blo t 5 template (2JMH) and that have been carried is: 10 for Der f 5 (derf5_monomer$_{1\rightarrow10}$[2JMH]), 10 for Sui m 5 (suim5_monomer$_{1\rightarrow10}$[2JMH]), 9 for Lep d 5 (lepd5_monomer$_{1\rightarrow9}$[2JMH]), 8 for Der f 21 (derf21_monomer$_{1\rightarrow8}$[2JMH]) and 9 for Der p 21 (derp21_monomer$_{1\rightarrow9}$[2JMH]).

Protein-Protein Docking Simulations

The ClusPro2.0 server (http://ClusPro.bu.edu/) (19), one of the top performers at CAPRI (Critical Assessment of Predicted Interactions) round 13-19 (20), was used to predict the possible dimeric structure of the different allergens. We submitted to ClusPro all the models of the monomers obtained with the 3MQ1_A template and selected during the comparative modelling stage (see previous section). ClusPro selects the 1000 best scoring solutions and then clusters them according to root mean square deviation (rmsd) considerations. Each cluster is characterized by its number of members, the ClusPro score of the center of the cluster and the lowest ClusPro score found in the cluster. We used the balanced ClusPro score (19). For each monomeric input, we kept the dimeric solution presenting the lowest ClusPro score.

We obtained the following structures: blot5_dimer$_{1\rightarrow8}$, derf5_dimer$_{1\rightarrow5}$, suim5_dimer$_{1\rightarrow7}$, lepd5_dimer$_{1\rightarrow6}$, derf21_dimer$_{1\rightarrow7}$, derp21_dimer$_{1\rightarrow8}$. These structures contain two monomers with kinked N-terminal helices.

Analysis of the Predicted Structure of the Dimers: Evaluation of the Root Mean Square Deviation (rmsd)

To establish the structural similarity between two structures, U3BEST (21) was used to superimpose all the C^{α} atoms and to compute the total rmsd deviation between the C^{α}'s after superimposition.

Determination of the Interactions at the Interface of the Dimers

The Protein Interaction Calculator (PIC) (22) was used to detect the various interactions at the interface of the dimers. Residues involved in hydrophobic contacts, salt bridges and in hydrogen bonds were identified. The percentage of common interactions of type *i* (hydrophobic contact, salt bridge or hydrogen bond) at the interfaces of the dimers *X* and *Y*, $P^{i}_{X/Y}$, has been computed according to the following equation:

$$P^{i}_{X/Y} = \frac{N^{i}_{X/Y}}{N^{i}_{Y}} \cdot 100\% \qquad [2]$$

where $N^{i}_{X/Y}$ is the number of common interactions of type *i* at the interfaces of the dimers *X* and *Y* and N^{i}_{Y} is the number of interactions of type *i* at the interface of the dimer *Y*.

Identification of Cavities

The Computed Atlas of Surface Topography of Proteins with structural and topographical mapping of functionally annotated residues "CASTp" (http://cast.engr.uic.edu) (23) was used to identify the cavities that are present in a given structure.

Results and Discussion

Analysis of the Multiple Sequence Alignment

The sequences of the allergens we studied are largely identical and/or similar (see Figure 1 and Table I): the identity is larger than 40%, except for the couples Der p 5/

Lep d 5, Der p 5/Der f 21, Der p 5/Der p 21, Der f 5/Der f 21 and Der f 5/Der p 21. Note that Der p 5 and Der f 5 present a very high sequence identity and similarity.

The residues that interact at the interface between chains A and B of the experimental structure of Der p 5 have been identified with the PIC program (22) and

PDB numbering (3MQ1)																									
Alignment numbering	1				5					10					15					20					25
Der p 5	H	D	Y	Q	N	E	F	D	F	L	L	M	E	R	I	H	E	Q	I	K	K	G	E	L	A
Blo t 5	D	D	F	R	N	E	F	D	H	L	L	I	E	Q	A	N	H	A	I	E	K	G	E	H	Q
Der f 5	H	D	Y	Q	N	E	F	D	F	L	L	M	Q	R	I	H	E	Q	M	R	K	G	E	E	A
Sui m 5	N	D	F	R	H	E	F	D	Y	L	L	M	K	T	A	E	H	N	M	E	R	G	E	A	M
Lep d 5	D	D	F	R	N	E	F	D	R	L	L	I	H	M	T	E	E	Q	F	A	K	L	E	Q	A
Der p 21	D	E	W	R	M	A	F	D	R	L	M	M	E	E	L	E	T	K	I	D	Q	V	E	K	G
Der f 21	D	K	W	R	N	A	F	D	R	M	L	M	E	E	F	E	E	K	I	D	Q	I	E	H	G

PDB numbering (3MQ1)	50				55					60					65					70					
Alignment numbering					30					35					40					45					50
Der p 5	L	F	Y	L	Q	E	Q	I	N	H	F	E	E	K	P	T	K	E	M	K	D	K	I	V	A
Blo t 5	L	L	Y	L	Q	H	Q	L	D	E	L	N	E	N	K	S	K	E	L	Q	E	K	I	I	R
Der f 5	L	L	H	L	Q	H	Q	I	N	T	F	E	E	N	P	T	K	E	M	K	E	Q	I	L	G
Sui m 5	L	L	A	L	T	E	Q	I	A	H	L	E	Q	S	K	N	K	E	E	K	E	K	I	V	R
Lep d 5	L	A	H	L	S	H	Q	V	T	E	L	E	K	S	K	S	K	E	L	K	A	Q	I	L	R
Der p 21	L	L	H	L	S	E	Q	Y	K	E	L	E	K	T	K	S	K	E	L	K	E	Q	I	L	R
Der f 21	L	L	M	L	S	E	Q	Y	K	E	L	E	K	T	K	S	K	E	L	K	E	Q	I	L	R

PDB numbering (3MQ1)	75				80					85					90					95					
Alignment numbering					55					60					65					70					75
Der p 5	E	M	D	T	I	I	A	M	I	D	G	V	R	G	V	L	D	R	L	M	Q	R	K	D	L
Blo t 5	E	L	D	V	V	C	A	M	I	E	G	A	Q	G	A	L	E	R	E	L	K	R	T	D	L
Der f 5	E	M	D	T	I	I	A	L	I	D	G	V	R	G	V	L	N	R	L	M	K	R	T	D	L
Sui m 5	E	L	E	T	I	I	A	L	I	S	G	S	H	D	V	L	E	R	E	L	K	R	T	D	L
Lep d 5	E	I	S	I	G	L	D	F	I	D	S	A	K	G	H	F	E	R	E	L	K	R	A	D	L
Der p 21	E	L	T	I	G	E	N	F	M	K	G	A	L	K	F	F	E	M	E	A	K	R	T	D	L
Der f 21	E	L	T	I	A	E	N	Y	L	R	G	A	L	K	F	M	Q	Q	E	A	K	R	T	D	L

PDB numbering (3MQ1)	100				105					110					115					120					
Alignment numbering					80					85					90					95					100
Der p 5	D	I	F	E	Q	Y	N	L	E	M	A	K	K	S	G	D	I	L	E	R	D	L	K	K	E
Blo t 5	N	I	L	E	R	F	N	Y	E	E	A	Q	T	L	S	K	I	L	L	K	D	L	K	E	T
Der f 5	D	I	F	E	R	Y	N	V	E	I	A	L	K	S	N	E	I	L	E	R	D	L	K	K	E
Sui m 5	D	I	L	E	R	Y	N	F	E	S	A	L	K	I	G	A	I	L	V	R	D	L	K	A	A
Lep d 5	N	L	A	E	K	F	N	F	E	S	A	L	S	T	G	A	V	L	H	K	D	L	T	A	L
Der p 21	N	M	F	E	R	Y	N	Y	E	F	A	L	E	S	I	K	L	L	I	K	K	L	D	E	L
Der f 21	N	M	F	E	R	Y	N	F	E	T	A	V	S	T	I	E	I	L	V	K	D	L	A	E	L

PDB numbering (3MQ1)	125				130					
Alignment numbering					105				110	
Der p 5	E	A	R	V	K	K	I	E	V	-	-	-	-
Blo t 5	E	Q	K	V	K	D	I	Q	T	Q	-	-	-
Der f 5	E	Q	R	V	K	K	I	E	V	-	-	-	-
Sui m 5	A	T	K	V	K	A	I	E	T	K	-	-	-
Lep d 5	E	A	K	V	K	A	I	N	V	H	A	-	-
Der p 21	A	K	K	V	K	A	V	N	P	D	E	Y	Y
Der f 21	A	K	K	V	K	A	V	K	S	D	D	-	-

Figure 1: Multiple sequence alignment of the allergens studied. This alignment has been performed with ClustalW (13). The first row contains the residue numbering in the Der p 5 structure (3MQ1). The next row corresponds to the alignment numbering. We used the PIC program (22) to identify the residues that interact at the interface of the Der p 5 dimer (chains A and B of 3MQ1). These residues are coloured in yellow in the Der p 5 row. The residues of chains A and B that interact in 3MQ1 and that are conserved in the other allergens are in yellow, those for which the hydrophobic character, the acidic character or the polar character is conserved are in green, pink and blue, respectively.

are presented with a yellow background on Figure 1 (Der p 5 row). The sequence identities computed between Der p 5 and the other allergens for these residues are given in Table II; they are larger that those obtained on the whole sequences, except for Lep d 5. Interestingly the sequence similarity between Der p 5 and Der f 5 on the interface residues is about 95% (see Table II).

Lep d 5, Der p 21 and Der f 21 do not contain a glycine residue at the alignment position 22 (45 in the PDB numbering of Der p 5, see Figure 1) and show a low sequence identity at the positions involved in the interface interactions in Der p 5 dimer. Since this glycine is suspected to play an important role in the formation of a kink in the N-terminal helix, allowing an entangled dimeric structure (4), it suggests that Lep d 5, Der p 21 and Der f 21 do not adopt a dimeric structure similar to that of Der p 5.

Table I
Sequence identity and similarity between the allergens studied.

	Blo t 5	Der p 5	Der f 5	Sui m 5	Lep d 5	Der f 21	Der p 21
Blo t 5	/	44% (107)	47% (107)	53% (107)	48% (110)	44% (109)	44% (107)
Der p 5	71% (107)	/	77% (109)	51% (109)	37% (107)	33% (106)	30% (106)
Der f 5	74% (107)	90% (109)	/	53% (109)	42% (107)	37% (106)	35% (106)
Sui m 5	74% (107)	68% (109)	66% (109)	/	46% (107)	43% (105)	41% (108)
Lep d 5	64% (110)	59% (107)	61% (107)	64% (107)	/	50% (109)	50% (107)
Der f 21	67% (109)	57% (106)	60% (106)	66% (107)	71% (109)	/	74% (111)
Der p 21	64% (107)	55% (106)	59% (106)	63% (108)	69% (107)	87% (111)	/

Sequence identity (upper triangle of the matrix) and similarity (lower triangle of the matrix) between the allergens studied. The length of the alignment is given between parentheses. The software used to perform the alignments, the scoring matrix used and the threshold to evaluate the similarity are given in Methods.

According to Mueller *et al.* (4), the valines at positions 62 and 65 (alignment numbering; positions 85 and 88 in the PDB numbering of Der p 5, see Figure 1) play an important role in the stability of the dimeric structure of Der p 5. Both valines are conserved solely in Der f 5.

On the contrary to what we described for Lep d 5, Der f 21 and Der p 21, a first analysis of the sequence alignment and of the sequence identity between Der p 5 and the other allergens suggest that Blo t 5, Der f 5 and Sui m 5 could adopt a dimeric structure similar to Der p 5, with a large hydrophobic cavity between both monomers, the most extensive sequence conservation being found between Der p 5 and Der f 5.

We then generated models of the monomeric structure of each allergen, and finally models of their dimeric structures.

Comparative Modelling of the Structure of the Monomers

The 3D structure of Blo t 5 (PDB code 2JMH) and Der p 5 (PDB code 3MQ1) having been resolved experimentally, we used both PDB's 2JMH and chain A of 3MQ1 to create a model of the structure of Der f 5, Sui m 5, Lep d 5, Der f 21 and Der p 21 by comparative modelling. We also computed a model presenting a kinked N-terminal helix for Blo t 5 using chain A of 3MQ1 as template.

The sequence similarity between the selected allergens being large enough (see Table I), we used a comparative modelling approach to predict the structure of the monomers of the allergens. We applied MODELLER9v4 (15) to build the three-dimensional models of the allergens. A first set of models was obtained with the structure of Blo t 5 (2JMH) as template. These models present a straight N-terminal helix and their energy has been evaluated with Anolea and Gromos (see Methods). The selected models, according to our energy criterion (see Methods), are given in Table SIa (see Supplementary material).

A second set of models was computed using the structure of chain A of Der p 5 (3MQ1_A) as template. The N-terminal helix is kinked in these models. The energy of the selected models is given in Table SIb (see Supplementary material).

Table III provides the average energy, computed with Gromos and Anolea, of the selected models and the lowest energy obtained with each template. The comparison of these values indicates, for each allergen, if the structure presenting a straight N-terminal helix (2JMH template) is predicted with a better score than that with a kinked helix (3MQ1_A template). In the case of Der f 5, the structure modelled with the 3MQ1_A template is clearly preferred compared to that obtained with the other

Table II
Sequence identity and similarity between Der p 5 and the other allergen, computed on the positions involved in the interactions at the dimeric interface of 3MQ1_AB.

	Identity (%)	Similarity (%)
Blo t 5	58	70
Der f 5	79	95
Sui m 5	58	70
Lep d 5	33	50
Der f 21	41	58
Der p 21	33	58

The sequence identity and similarity between Der p 5 and the other allergen have been computed on the 24 positions involved in the interactions at the dimeric interface of 3MQ1_AB (see Figure 1). The software used to perform the alignments, the scoring matrix and the threshold to evaluate the similarity are given in Methods.

Table III
Energy of the monomeric models computed with 2JMH or 3MQ1_A as template.

	Lowest energy anolea		Lowest energy gromos		Average energy anolea		Average energy gromos	
Template	3MQ1_A	2JMH	3MQ1_A	2JMH	3MQ1_A	2JMH	3MQ1_A	2JMH
Der f 5	**−232.0**	−190.0	**−11283.3**	−9805.5	**−207.2**	−173.3	**−10008.0**	−8427.9
Δ	−42.0		−1477.8		−33.9		−1580.1	
Sui m 5	**−186.0**	−111.0	−8609.9	**−9078.4**	**−151.0**	−79.3	**−8036.2**	−7662.0
Δ	−75		468.5		−71.7		−374.2	
Lep d 5	−232.0	−232.0	**−8569.8**	−8320.8	**−194.3**	−186.4	−7452.4	**−7748.1**
Δ	0.0		−249.0		−7.9		295.7	
Der p 21	**−189.0**	−109.0	−8627.2	**−8963.9**	**−144.8**	−72.2	**−8075.4**	−7778.1
Δ	−80.0		336.7		−72.6		−297.3	
Der f 21	**−188.0**	−173.0	−7215.5	**−7297.1**	**−133.0**	−101.0	**−6731.0**	−5790.9
Δ	−15.0		81.6		−32		−940.1	
Blo t 5	−235.0	/	−10256.0	/	−154.9	/	−9188.5	/

The energies are computed with Anolea and with Gromos. The model with the lowest energy and the average energy computed on all the selected models are provided. Δ is the difference between the energy of the models computed with the 3MQ1_A and the 2JMH templates, respectively.

template. The results are less clear-cut for the models of Sui m 5, Lep d 5, Der p 21 and Der f 21: we reach different conclusions according to the criterion used to define which template leads to the best model (lowest energy or average energy computed with Anolea or Gromos). Moreover, the Gromos energy difference between the models computed with the two templates is significantly larger in the case of Der f 5 compared to the other allergens (see row "Δ" in Table III). This result suggests that a kinked N-terminal helix would be more probable in the Der f 5 monomer than in the other allergens studied. Note that Lep d 5, Der p 21 and Der f 21 do not present a Gly residue at position 22 of the sequence alignment. The presence of another residue at this position could hinder the formation of a kink in the helix.

We also used 3MQ1_A as template to obtain a kinked structure for the Blo t 5 monomer. Interestingly, the energy of such a structure is quite low, not as much as the energy obtained for the Der f 5 models, but lower than those computed for Sui m 5, Lep d 5, Der p 21 and Der f 21 (see Table III).

Reliability of Cluspro to Predict the Structure of the Dimers

We used the ClusPro2.0 server (19) to predict the structure of the dimers. To test the ability of ClusPro to predict the dimeric structure of these allergens, we first used it to compute the structure of the Der p 5 dimer and compared this predicted structure to the X-ray structure of the dimer (chains A and B of the 3MQ1 PDB, 3MQ1_AB).

The chain A of 3MQ1 was submitted to the ClusPro server and 20 solutions were returned. We selected the structure with the lowest ClusPro score, and we called this structure derp5_dimer[3MQ1_A].

The root mean square deviation (rmsd) after superimposition of the C^{α} atoms between 3MQ1_AB and derp5_dimer[3MQ1_A] (see Methods) is equal to 1.9 Å, which reflects that the predicted structure is close to the experimentally resolved one.

The interactions at the interface of 3MQ1_AB and of derp5_dimer[3MQ1_A] were characterized with the PIC program (22), and compared (see Figure 2): 36 interactions were found in the experimental structure (24 interactions between hydrophobic

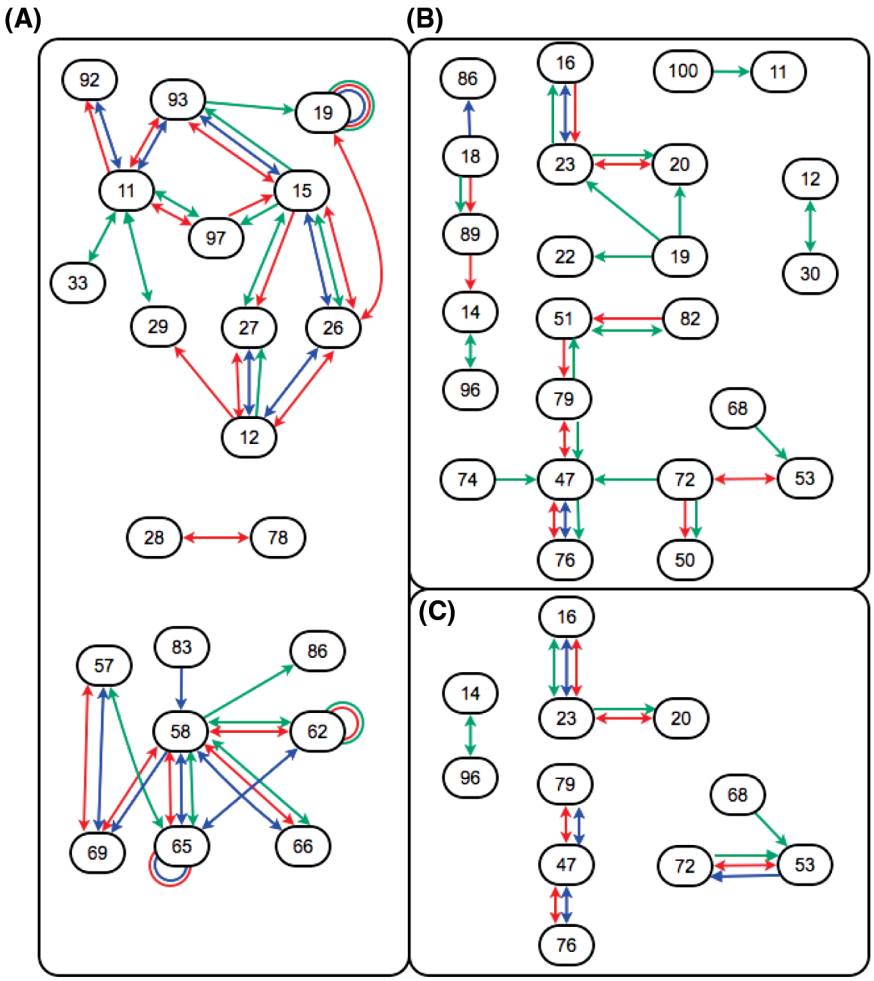

Figure 2: Residues that interact at the interface of the dimers. The interface interactions were detected with the PIC program (22). The residue positions are numbered according to the sequence alignment numbering (see Figure 1). Positions that interact in the experimental structure of Der p 5 (3MQ1_AB), in the dimeric structure of Der p 5 predicted by ClusPro (derp5_dimer[3MQ1_A]) and in the predicted dimeric structure of Der f 5 (derf5_dimer$_5$) are connected by a blue, red and green arrow, respectively. The dimers are not perfectly symmetric and the detected interactions between chains A and B are not systematically the same as those between chains B and A. The arrows of the connections are oriented from the interacting residue in chain A toward the residue in chain B. (**A**) Hydrophobic contacts between residues. (**B**) Hydrogen bonds between residues. (**C**) Salt bridges between residues.

residues, 5 hydrogen bonds and 7 ionic interactions) and 57 interactions were found in the predicted structure (33 interactions between hydrophobic residues, 14 hydrogen bonds and 10 ionic interactions). Table IV indicates that 83% of the interface interactions found in 3MQ1_AB are common with derp5_dimer[3MQ1_A], and that 53% of the interface interactions found in derp5_dimer[3MQ1_A] are common with 3MQ1_AB. Table IV reveals also that these percentages are larger when they are computed only on the hydrophobic contacts or on the salt bridges, and lower on the hydrogen bonds. There are more hydrogen bonds at the interface of derp5_dimer[3MQ1_A] than in the experimental structure, but most of the additional hydrogen bonds are located in the same region as the native ones (see Figure 3A).

As discussed above, Mueller *et al.* (4) suggested that the valines at positions 62 and 65 (alignment numbering) could play an important role in the stability of the dimeric structure of Der p 5 and could be responsible for the difference in the dimerization propensity between Der p 5 and Blo t 5. These hydrophobic contacts are not identified with PIC in derp5_dimer[3MQ1_A] (see Figure 2). However, an analysis of the structure shows that the residues 62 and 65 present a similar orientation in 3MQ1_AB and in derp5_dimer[3MQ1_A] (see Figure 3B).

We then submitted 3MQ1_AB and derp5_dimer[3MQ1_A] to the CASTp program (23) in order to characterize the cavity that could be involved in ligand binding. The large cavity at the interface of both monomers presents a volume of 3051 Å^3 in 3MQ1_AB and of 2914 Å^3 in derp_dimer[3MQ1_A]. Moreover, 83% of the

Table IV
Percentage of common interface interactions between 3MQ1_AB, derp5_dimer[3MQ1_A] and derf5_dimer$_5$.

	3MQ1_AB			derp5_dimer[3MQ1_A]			derf5_dimer$_5$		
	$P^{hydrophob}_{3MQ1/X}$	$P^{Hbond}_{3MQ1/X}$	$P^{salt\text{-}bridge}_{3MQ1/X}$	$P^{hydrophob}_{derp5/X}$	$P^{Hbond}_{derp5/X}$	$P^{salt\text{-}bridge}_{derp5/X}$	$P^{hydrophob}_{derf5/X}$	$P^{Hbond}_{derf5/X}$	$P^{salt\text{-}bridge}_{derf5/X}$
3MQ1_AB		/		83%	60%	100%	38%	40%	43%
					$P_{derp5/X}$ = 83%			$P_{derf5/X}$ = 39%	
derp5_dimer[3MQ1_A]	61%	21%	70%		/		45%	43%	40%
		$P_{3MQ1/X}$ = 53%						$P_{derf5/X}$ = 44%	
derf5_dimer$_5$	36%	10%	43%	60%	30%	57%		/	
		$P_{3MQ1/X}$ = 27%			$P_{derp5/X}$ = 48%				

Percentage of common interface interactions between the experimental dimeric structure of Der p 5 (3MQ1_AB), the dimeric structure of Der p 5 predicted by ClusPro (derp5_dimer[3MQ1_A] and the best dimeric structure of Der f 5 predicted by ClusPro (derf5_dimer$_5$). See equation 2 for the definition of these percentages (see Methods). X is the structure corresponding to the considered line in the Table. The interactions are given in Figure 2. $P^{hydrophob}$, P^{Hbond} and $P^{salt\text{-}bridge}$ are computed on the residues involved in hydrophobic effects, hydrogen bonds and salt bridges, respectively.

residues involved in the cavity of 3MQ1_AB are also present in the cavity of derp5_dimer[3MQ1_A].

All these results suggest that ClusPro is able to reliably predict the structure of the dimers of the allergens studied here and that the model with the lowest ClusPro score corresponds to the prediction with the lowest rmsd compared to the native structure.

Predicted Dimeric Structure of the Allergens and Analysis of their Properties

The crystal structure of the Der p 5 dimer contains a large hydrophobic cavity that could be involved in ligand binding (4). To test whether the various allergens studied could adopt a dimeric structure similar to Der p 5 and show a hydrophobic cavity, we submitted to ClusPro all the selected monomeric models that have a kinked helix (see Table SIb in Supplementary Material); these monomers were modelled with 3MQ1_A as template. We obtained several possible solutions with ClusPro, and the solution with the lowest ClusPro score was selected. Indeed, the results obtained in the previous section suggested that the ClusPro score is able to discriminate our models and that the structure with the lowest ClusPro score should be the most similar to the native structure. We finally obtained eight homodimers for Blo t 5, five for Der f 5, seven for Sui m 5, six for Lep d 5, eight for Der p 21 and seven for Der f 21. The ClusPro scores of the best homodimeric solutions are given in Table V, as well as their rmsd with 3MQ1_AB; see Table SII (Supplementary Material) for the scores and rmsd's of all the predicted homodimers.

Der f 5 is the only allergen that presents a predicted dimeric structure similar to that of Der p 5 (see Table V, derf5_dimer$_5$). The rmsd after superimposition between Der p 5 (3MQ1_AB) and derf5_dimer$_5$ is equal to 3.0 Å; 48% of the interactions identified at the interface of derf5_dimer$_5$ are common to those found in

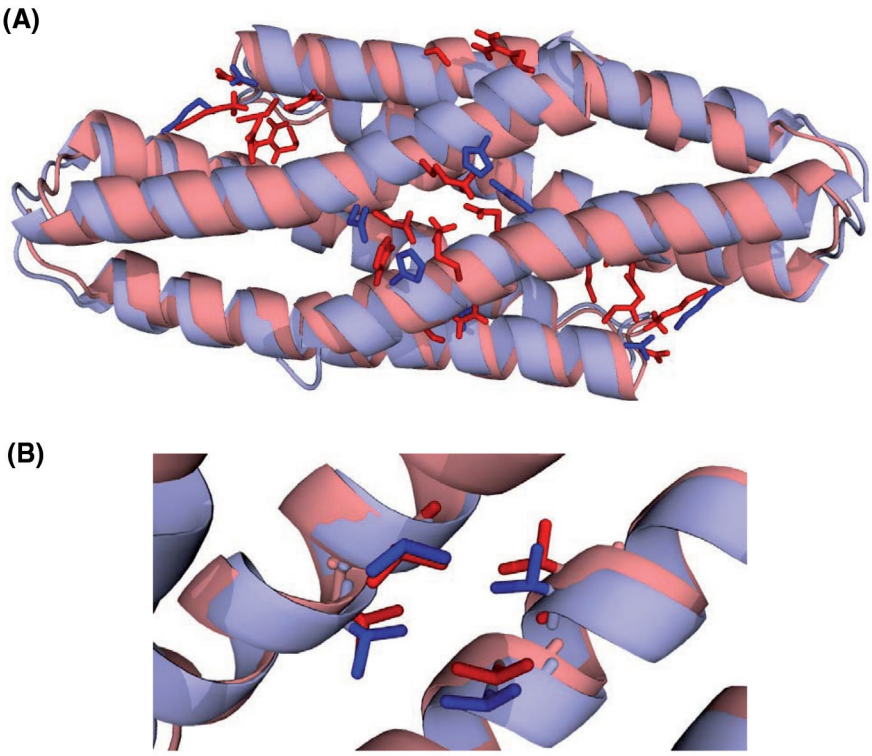

Figure 3: Superimposition between the experimental dimeric structure of Der p 5 (chains A and B of the 3MQ1 PDB, in light blue) and the dimeric structure of Der p 5 predicted with ClusPro (in salmon); the structures are referred as 3MQ1_AB and derp5_dimer[3MQ1_A], respectively. The rmsd after superimposition is equal to 1.9 Å. The residues that are in interaction at the interface of each monomer have been detected with the PIC program (22). (**A**) The residues that are involved in hydrogen bonds at the interface of 3MQ1_AB and of derp5_dimer[3MQ1_A] are in blue and in red, respectively. (**B**) The valines number 62 and 65 (alignment numbering, see Figure 1) that interact at the interface of 3MQ1_AB are in blue and in red in 3MQ1_AB and in derp5_dimer[3MQ1_A], respectively.

Table V
Best models of the dimeric structure of the different allergens.

	ClusPro score	Rmsd (Å)
derp5_ dimer[3MQ1_A]	−1188	1.9
blot5_dimer$_6$	−1152	>10
derf5_dimer$_5$	−1353	3.0
suim5_dimer$_4$	−1168	>10
lepd5_dimer$_3$	−1190	>10
derp21_dimer$_6$	−1266	>10
derf21_dimer$_2$	−1252	>10
derp5_V62AV65A	−1053	2.1
derp5_L69E	−1063	1.8
derp5_ V62AV65AL69E	−930	2.1
derp5_all	−869	>10
derf5_V62AV65A	−1218	4.2
derf5_L69E	−1169	2.9
derf5_ V62AV65AL69E	−1039	3.1
blot5_A62VA65V	−1266	>10
blot5_E69L	−1289	>10
blot5_ A62VA65VE69L	−1401	>10
blot5_all	−1090	>10
blot5_dimer2_all	−1057	4.0

The dimeric structures have been computed with ClusPro, using as monomer the models obtained with the 3MQ1_A template. These monomers present a kinked helix, as observed in the experimental dimeric structure of Der p 5. The dimer number 6, for instance, corresponds to the best ClusPro prediction obtained with the monomer number 6 (see Table SIb). The rmsd is computed between the model and the 3MQ1_AB structure.

derp5_dimer[3MQ1_A] (see Figure 2 and Table IV). We showed in the previous sections that slight differences in side-chains orientation can influence the interface interactions detected by PIC, leading to a larger number of interactions identified in derp5_dimer[3MQ1_A] than in 3MQ1_AB. For this reason, we estimated that it was sounder to compare the interface interactions of derf5_dimer$_5$ to those found in derp5_ dimer[3MQ1_A]. The volume of the cavity obtained by the assembly of both monomers in derf5_dimer$_5$ is equal to 1979 Å^3.

Mueller *et al.* (4) discussed the possible origin of the difference in the dimerization propensity between Der p 5 and Blo t 5. They identified the sequence region GVRGV (alignment numbering 61-65), which is GAQGA in Blo t 5, and they suggested that the hydrophobic contacts between the valines of this region could be responsible for the different dimerization propensity of Blo t 5 and Der p 5. Interestingly, the best dimeric model of Der f 5, derf5_dimer5, is similar to the structure of the Der p 5 dimer. This is in agreement with the large sequence identity between both proteins, and with the sequence conservation of the residues involved in the interface interactions (see Tables I and II). Moreover, the sequence region 61-65 of the alignment (GVRGV in Der p 5) is conserved in Der f 5. The alignment position 69 could also play a role in the difference between Der p 5, Blo t 5 and Der f 5: a Leu residue that makes hydrophobic contacts (see Figure 2) is present in Der p 5 and in Der f 5, whereas a negative charge (Glu) is found in Blo t 5. We modelled mutated Blo t 5, Der p 5 and Der f 5 proteins and we submitted them to ClusPro to test the assumption that positions 62, 65 and 69 along the alignment play an important role in the dimeric structure of Der p 5. More precisely, we chose the monomers that led to the best dimeric structure (3MQ1_A, blot5_ monomer$_6$[3MQ1_A] and derf5_monomer$_5$[3MQ1_A]), introduced in Der p 5 and

in Der f 5 the residues that are found in Blo t 5, and vice versa, and we submitted these mutated structures to ClusPro. We tested the following mutations: V62A-V65A, L69E, V62A-V65A-L69E in Der p 5 and in Der f 5 (derp5_V62AV65A, derf5_V62AV65A, derp5_L69E, derf5_L69E, derp5_V62AV65AL69E, derf5_V62AV65AL69E, respectively), A62V-A65V, E69L, A62V-A65V-E69L in Blo t 5 (blot5_A62VA65V, blot5_E69L, blot5_A62VA65VE69L, respectively). We also identified and mutated all the positions that are involved in the interactions at the interface of the Der p 5 dimer and that are not conserved in Blo t 5, and vice versa: M12I-R14Q-I15A-H16N-E17H-Q18A-K20E-L24H-H35E-A50R-T54V-I55V-V62A-V65A-L69E-D76N-L83Y-M85E-S89L in Der p 5 (derp5_all), and I12M-Q14R-A15I-N16H-H17E-A18Q-E20K-H24L-E35H-R50A-V54T-V55I-A62V-A65V-E69L-N76D-Y83L-E85M-L89S in Blo t 5 (blot5_all).

The results are given in Table V. The mutations V62A, V65A and L69E in Der p 5 and in Der f 5 led to a predicted dimer that is still close to the experimental structure of Der p 5: the rmsd is lower than or equal to 3.1 Å. However, the ClusPro score increases, suggesting a less stable dimer. The mutation of the interface residues in Der p 5 to the Blo t 5 residues (derp5_all) gives a dimeric structure which is no more native-like and with a larger ClusPro score. The residues at positions 62, 65 and 69 seem thus to play a role in the stability of the dimer, but, according to our simulations, their mutation is not sufficient to impair the formation of the native dimeric structure.

The mutation of the Blo t 5 sequence (A62V, A65V and E69L) does not promote the formation of a dimeric structure that is similar to Der p 5: even if the ClusPro score decreases, the rmsd is larger than 10 Å (see Table V). Blot5_all contains all the residues that are involved in the interface interactions of Der p 5. We were thus surprised that the dimeric structure obtained with blot5_all still present a very large rmsd compared to the Der p 5 native dimeric structure (see Table V). The analysis of the structures revealed that the N-terminal helix is less kinked in the Blo t 5 monomer (blot5_monomer$_6$[3MQ1_A]), which complicates the formation of an entangled dimer. The helix is more kinked in blot5_monomer$_2$[3MQ1_A] (see Table SIb in supplementary material). We introduced all the mutations listed above (I12M-Q14R-A15I-N16H-H17E-A18Q-E20K-H24L-E35H-R50A-V54T-V55I-A62V-A65V-E69L-N76D-Y83L-E85M-L89S) in this structure, and we submitted it to ClusPro. The best dimeric structure (blot5_dimer2_all) presents an rmsd of 4.0 Å with the Der p 5 experimental structure (see Table V). It is thus necessary to mutate a large number of sequence positions to induce an entangled dimeric structure for Blo t 5.

Finally, in the case of Suim 5, Lep d 5 and Der p 21, we found dimeric models close to the Der p 5 experimental structure, but these predictions present a larger ClusPro score (see Table SII in supplementary material). This kind of dimer is thus less probable for these proteins. Moreover, Lep d 5 and Der p 21 do not contain the glycine at position 22 of the alignment. This glycine is assumed to be responsible for the kink in the N-terminal helix, which is necessary to form the entangled dimeric structure.

Conclusion

Equivalents of Blo t 5 and Der p 5 that share 44% sequence identity were isolated from different house dust mite species. The published structure of Blo t 5 revealed a monomeric three helical bundle (9,10), whereas the structure of Der p 5 is a three helical bundle with a kinked N-terminal helix that assembles in an entangled dimer with a large hydrophobic cavity (4). This cavity could be involved in the binding of some hydrophobic ligands, like LPS, that could be responsible for the shift of the immune response from tolerance to allergic inflammation (4, 24). We used molecular modelling approaches to analyze the possibility that other allergens of families 5 and 21 of house dust mites (Der f 5, Sui m 5, Lep d 5, Der p 21 and Der f 21) can adopt a dimeric structure similar to Der p 5, with a large hydrophobic cavity.

The modelling of the monomeric structure of the allergens with the structures of Der p 5 (kinked N-terminal helix) and Blo t 5 (straight N-terminal helix) as templates suggests that Der f 5 has a preference for the structure with a kinked helix (see Table III), whereas the conclusion is less clear-cut for the other allergens.

The kinked monomers were submitted to the ClusPro server in view of predicting the possible dimeric structure of these allergens. This kind of server is unable to predict if a protein will or will not form a dimer but allowed us to look whether a potential dimer would have the same characteristics than the one of Der p 5 dimer and would present a low ClusPro score. We predicted that Der f 5 could form such a dimer, with a cavity about 2000 $Å^3$.

Mueller *et al.* (4) also suggested that the differences between Blo t 5 and Der p 5 at positions 62 and 65 of the sequence alignment (see Figure 1) could explain the fact that Blo t 5 is monomeric and Der p 5 dimeric. The sequence alignment reveals that the neighbouring position 69 could also participate in the difference between both proteins: a Leu is present in Der p 5 and a Glu in Blo t 5 (see Figure 1). Moreover these three sequence positions are conserved between Der p 5 and Der f 5. We mutated the Der p 5 residues at these positions into the Blo t 5 residues, and vice versa. The results of the ClusPro prediction suggest that these residues could play a role in the stability of the dimer. But mutating them is not sufficient to avoid the formation of an entangled dimeric structure for Der p 5 and to promote it for Blo t 5. We also identified the sequence positions that are involved in the interactions at the interface of the Der p 5 dimer and we analyzed their conservation in Blo t 5 and in Der f 5. In Der p 5 we mutated all the positions that are not conserved, and introduced the residues found in Blo t 5, and vice versa. The rmsd between the native structure of Der p 5 and the predicted dimeric structure of the mutated Der p 5, which contains Blo t 5-like interface residues, is larger than 10 Å. This result suggests that it is necessary to mutate a large number of residues to lose the entangled dimeric structure of Der p 5. Interestingly, doing the reverse operation in Blo t 5 can lead to an entangled dimeric structure with a cavity of 1585 $Å^3$.

Our results suggest several experiments. Only Der p 5 and Der f 5 should present a similar dimeric structure, with the formation of a large cavity that could accommodate some hydrophobic ligands. The other allergens studied here – Blo t 5, Sui m 5, Lep d 5, Der p 21 and Der f 21 – should not present such a propensity. It could thus be tested experimentally whether Der f 5 adopts an entangled dimeric structure with a large hydrophobic cavity that is similar to Der p 5. Second, we identified mutations that could promote or hinder the dimerization propensity of Der p 5, Der f 5 and Blo t 5. The oligomerization of Der p 5 being concentration-dependent (4), mutations that decrease the stability of the dimer should increase the concentration needed to observe these dimers. Moreover, characterizing the mutations suggested by our study could help to understand if the formation of an hydrophobic cavity through the dimerization of Der p 5 plays an important role in the induction of an allergic response, and if Der p 5 and Der f 5 behave similarly.

Finally, it would be surprising if the allergens from family 5 induce an allergic response through different mechanisms. We showed that only Der f 5 is prone to form a dimer similar to Der p 5, and that we do not obtain such a dimer with Blo t 5 even if a modelled Blo t 5 structure with a kinked N-terminal helix has a quite low calculated energy. Therefore, we do not think that the dimerization process in Der p 5 and maybe in Der f 5 is involved by itself in the induction of the allergic response. However, the opening of the structure through the formation of a kink in the N-terminal helix, even if this conformational change is transitory, could provoke the exposition of some hydrophobic amino acids, the binding of hydrophobic ligands and induce the allergic response. The peculiarity of Der p 5 and probably Der f 5 would be that the kinked structure could be stabilized through dimerization.

Table SI: score of the models computed with 2JMH (Blo t 5) and 3MQ1_A (Der p 5) as template.

Table SII: Score of the models of the dimeric structure of the different allergens. These can be obtained free of charge from the authors or can be purchased from Adenine Press for US $50.00.

Acknowledgment

The Belgian State Science Policy Office through an Interuniversity Attraction Poles Program (DYSCO) supported this work and the Commission Universitaire pour le Développement (CUD) through a PIC project. The authors thank Prof. Paul Jacobs for his advice and careful reading.

References

1. K. Meno, P. B. Thorsted, H. Ipsen, O. Kristensen, J. N. Larsen, M. D. Spangfort, M. Gajhede, and K. Lund, *J Immunol 175*, 3835-3845 (2005).
2. S. de Halleux, E. Stura, L. VanderElst, V. Carlier, M. Jacquemin, and J.-M. Saint-Remy. *J Allergy Clin Immunol 117*, 571-576 (2006).
3. G. A. Mueller, D. C. Benjamin, and G. S. Rule. *Biochemistry 37*, 12707-12714 (1998).
4. G. A. Mueller, R. A. Gosavi, J. M. Krahn, L. L. Edwards, M. J. Cuneo, J. Glesner, A. Pomés, M. D. Chapman, R. E. London, L. C. Pedersen. *J Biol Chem 285*, 25394-25401 (2010).
5. G. A. Mueller, L. L. Edwards, J. J. Aloor, M. B. Fessler, J. Glesner, A. Pomés, M. D. Chapman, R. E. London, and L. C. Pedersen. *J Allergy Clin Immunol 125*, 909-917. e4 (2010).
6. M. Chruszcz, M. D. Chapman, L. D. Vailes, E. A. Stura, J.-M. Saint-Remy, W. Minor, and A. Pomés. *J Mol Biol 386*, 520-530 (2009).
7. B. R. Johannessen, L. K. Skov, J. S. Kastrup, O. Kristensen, C. Bolwig, J. N. Larsen, M. Spangfort, K. Lund, and M. Gajhede. *FEBS Lett 579*, 1208-1212 (2005).
8. S. L. Chan, S. T. Ong, S. Y. Ong, F. T. Chew, and Y. K. Mok. *J Immunol 176*, 4852-4860 (2006).
9. M. T. Naik, C.-F. Chang, I.-C. Kuo, C. C.-H. Kung, F.-C. Yi, K.-Y. Chua, and T.-H. Huang. *Structure 16*, 125-136 (2008).
10. S. L. Chan, T. C. Ong, Y. F. Gao, Y. S. Tiong, D. Y. Wang, F. T. Chew, and Y. K. Mok. *J Immunol 181*, 2586-2596 (2008).
11. W. R. Thomas, B. J. Hales, and W. A. Smith. *Curr Allergy Asthma 5*, 388-393 (2005).
12. M. Weghofer, Y. Dall'Antonia, M. Grote, A. Stçcklinger, M. Kneidinger, N. Balic, M.-T. Krauth, E. Fernández-Caldas, W. R. Thomas, M. van Hage, S. Vieths, S. Spitzauer, F. Horak, D.-I. Svergun, P.-V. Konarev, P. Valent, J. Thalhamer, W. Keller, R. Valenta, and S. Vrtala. *Allergy 63*, 758-767 (2008).
13. M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallase, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins. *Bioinformatics 23*, 2947-2948 (2007).
14. S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, W. Miller, and D. J. Lipman. *Nucleic Acids Res 25*, 3389-3402 (1997).
15. A. Sali and T. L. Blundell. *J Mol Biol 234*, 779-815 (1993).
16. P. W. Rose, B. Beran, C. Bi, W. F. Bluhm, D. Dimitropoulos, D. S. Goodsell, A. Prlic, M. Quesada, G. B. Quinn, J. D. Westbrook, J. Young, B. Yukich, C. Zardecki, H. M. Berman, and P. E. Bourne. *Nucleic Acids Res 39*, D392-D401 (2011).
17. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen. *J Comput Chem 26*, 1701-1718 (2005).
18. F. Melo and E. Feytmans. *J Mol Biol 277*, 1141-1152 (1998).
19. D. Kozakov, R. Brenke, S. R. Comeau, and S. Vajda. *Proteins 65*, 392-406 (2006).
20. D. Kozakov, D. R. Hall, D. Beglov, R. Brenke, S. R. Comeau, Y. Shen, K. Li, J. Zheng, P. Vakili, I. C. Paschalidis, and S. Vajda. *Proteins 78*, 3124-3130 (2010).
21. W. Kabsch. *Acta Crystallogr Sect A 34*, 827-828 (1978).
22. K. G. Tina, R. Bhadra, and N. Srinivasan. *Nucleic Acids Res 35*, W473-476 (2007).
23. J. Dundas, Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, and J. Liang. *Nucleic Acids Res 34*, W116-118 (2006).
24. A. Jacquet. *Trends Mol Med*, In press (2011).

Date Received: May 16, 2011

Communicated by the Editor Ramaswamy H. Sarma