

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE  
SCIENTIFIQUE

**UNIVERSITE M'HAMED BOUGARA-BOUMERDES**

---



Faculté Des Sciences

## **Thèse de Doctorat**

Présentée par

**KHEMILI Souad**

Filière: Sciences Biologiques

Option : Biologie Moléculaire (Bioinformatique)

---

***Prédiction des épitopes B de l'allergène Der p 5 et étude in silico de la structure et de l'état oligomérique des allergènes d'acariens des familles 5 et 21***

---

**Devant le jury:**

HAMMADI	Mohamed	Prof	UMBB	Président
GILIS	Dimitri	Prof	ULB	Directeur de thèse
GANNA	Salima	MC/A	UMBB	Co-directrice
JACQUET	Alain	Prof	UBT	Examinateur
HAMIDECHI	Med Hafid	Prof	UMC	Examinateur
ABDERAHMANI	Ahmed	MC/A	USTHB	Examinateur
ROOMAN	Marianne	Prof	ULB	Examinatrice
JACOBS	Paul	Prof	ULB	Invité

Année Universitaire: 2012/2013

## *Dédicaces*

*Je dédie ce travail*

*à mes très chers parents,*

*mes chers frères Mohamed et Hichem*

*mes chères soeurs Radia et Amina,*

*et mes adorables petits neveux AbdErahmane et Ayoub,*

*à Tata Djawida , Tonton Kamel et leurs petits anges (Ramzi, Kamelia, Melina et Faten),*

*à mes chères Tata Fifi, Tata Fadila et Tante Malika qui ne m'ont jamais oublié avec leurs prières,*

*ainsi qu'à tous mes amis qui étaient toujours là pour moi...*

*A la mémoire de mes grands-parents, qui, depuis mon enfance espéraient me voir Docteur un jour! J'aurai aimé tant que vous soyez avec moi en ce moment...*

## Remerciements

*Au terme de ce travail de thèse, je saisis cette occasion pour exprimer mes sincères remerciements à toute personne ayant contribué, de près ou de loin, à la réalisation de ce travail.*

*Cette thèse est le fruit d'une collaboration étroite entre notre université (U.M.B.B.) et l'Université Libre de Bruxelles (U.L.B.). Je souhaite, tout d'abord, remercier la "Commission Universitaire pour le Développement" qui a financé ce projet de recherche et qui m'a permis d'avoir une meilleure formation en tant que chercheur et en tant qu'enseignante.*

*Mes vifs remerciements s'adressent également à tous mes responsables de l'U.M.B.B., Madame le Recteur de l'université de Boumerdès, Monsieur le Vice-recteur chargé des relations internationales, Monsieur le Vice-recteur chargé de la post-graduation, Monsieur le Doyen de la Faculté des Sciences et Monsieur le Chef de Département de biologie, qui ont efficacement contribué au bon déroulement de la réalisation de cette thèse.*

*Je tiens à remercier notre ex-Recteur de l'université de Boumerdès, Madame le Professeur Kesri R., le Recteur de l'Université de Bouira qui était ex-Doyen de la Faculté des Sciences, Monsieur le Professeur Badari K., ainsi que notre ex-Chef de Département de biologie, Madame le Docteur Sahir-Halouane F. pour tous leurs encouragements et toutes les décisions difficiles qu'ils ont pris à mon égard afin de mener à bien ce projet de coopération qui a abouti à l'élaboration de cette thèse. Vous trouverez ici l'expression de ma profonde reconnaissance ainsi que de mes hommages les plus respectueux.*

*Cher Monsieur Jacobs P. (Professeur à l'Université Libre de Bruxelles, Belgique), je t'ai toujours considéré comme un père pour moi. Tu as toujours su m'orienter et me guider avec tous tes conseils et tes critiques durant nos longues discussions. J'ai beaucoup appris avec toi et ce, que se soit sur le plan scientifique ou culturel. Je t'admire tellement que certaines personnes ont l'impression que c'est toi qui leur parle et non pas moi! Ta contribution est d'une grande ampleur pour la réussite de tous nos projets de coopération. Je te remercie pour tout et, surtout, d'avoir accepté de me faire l'honneur de siéger dans mon jury.*

*L'expression de ma haute reconnaissance s'adresse à mon promoteur Monsieur Gilis D. (Professeur à l'Université Libre de Bruxelles, Belgique): ta gentillesse, ta rigueur scientifique, ta simplicité et ta patience m'ont fascinées. Je te remercie pour ta disponibilité à mon égard et ton dévouement à chaque instant où j'ai sollicité ton aide. Tes orientations, tes conseils et tes suggestions m'ont été très utiles pour ma formation et ont permis à ce que ce travail de thèse soit agréable et très fructifiant. Je te serai reconnaissante toute ma vie.*

Mes vifs remerciements s'adressent également à ma co-promotrice, Madame Gana-Kebbouche S. (Maître de conférences "A" à l'U.M.B.B.). Tu as toujours su me guider, me montrer le bon chemin et me soutenir à surmonter toutes mes difficultés quelles qu'elles soient. C'était et c'est toujours très agréable, pour moi, de travailler à tes côtés.

Je tiens également à remercier Monsieur Hamadouche T. (Maître de conférences "B" à l'U.M.B.B.), d'avoir accepté de m'encadrer durant ces années et de faire partie de mon jury.

Mes vifs remerciements s'adressent aussi à Monsieur Hamadi M. (Professeur à l'U.M.B.B.) pour m'avoir fait l'honneur de présider mon jury de thèse.

J'exprime également ma grande gratitude à Monsieur Jacquet A. (Professeur à l'Université Chulalongkorn de Bangkok, Thaïlande), à Madame Rooman M. (Professeur à l'Université Libre de Bruxelles, Belgique), à Monsieur Mamidechi H. (Professeur à l'Université Mentouri de Constantine) et à Monsieur Abdérahmani A. (Maître de conférences "A" à l'Université des Sciences et de la technologie Houari Boumediene) pour l'honneur qu'ils m'ont fait en acceptant de participer à ce jury et d'examiner mon travail.

Mes remerciements s'adressent également à Madame Marianne Rooman et à Monsieur Philippe Bogaerts pour leur accueil chaleureux dans le service 3BIO. Je tiens à remercier tous les membres de ce service: Jean Marc, Marie, Jaro, Suchsía, Yves, Fabian, Alex, Nathalie, Danièle, Anne, Zakaria, Patrick et Niu, sans oublier Malik, Tomasz et Jean-Louis. Le caractère aimable avec lequel vous m'avez toujours reçu et accueilli m'a profondément touché. Je ne vous oublierai jamais ...

Mes plus profonds remerciements vont, surtout et tout particulièrement, à mes très chers parents qui m'ont soutenu, donné confiance et supporté tous mes caprices durant tout mon parcours. Je ne trouverai jamais assez de mots pour décrire ce que j'éprouve pour vous. J'ai eu le plaisir de faire toute cette carrière près de vous et vous voir si fiers de moi est ma plus grande réussite: que Dieu vous garde pour moi...

Enfin, une thèse est une période avec de bons moments. D'autres le sont moins et ce sont dans ces derniers que l'on reconnaît le véritable soutien. Je souhaite donc remercier tous mes proches, en particulier Zakaria, Sihem, Sadjia, Nesrine et Amina qui étaient toujours là pour me soutenir et m'encourager lors de la réalisation de cette thèse qui demeure la plus merveilleuse expérience humaine et scientifique que je porterai toujours dans mon cœur.

## Résumé:

Cette recherche s'inscrit dans le cadre de l'étude *in silico* (en utilisant des approches bioinformatiques et de modélisations moléculaires) des structures 3D des allergènes d'acariens de poussières de maisons des familles 5 et 21. Nous avons pour objectifs de définir les épitopes de Der p 5 et de fournir des informations quant à leur propension à dimériser. Cela contribue à l'étude de leurs mécanismes d'allergénicité et pourrait aboutir à la conception de variants hypoallergéniques pour l'immunothérapie. Quatre régions antigéniques ont été identifiées pour Der p 5: H34-L39 (épitope-Nt), H58-D69 (épitope II), L92-E107 (épitope I) et K129-V132 (épitope-Ct) (selon la numérotation du code PDB 3MQ1 de Der p 5). Par la suite, les résidus situés dans ces régions qui pourraient jouer un rôle crucial dans la liaison aux immunoglobulines E (IgE) ont été sélectionnés selon des critères que nous avons définis (nature des acides aminés et leur accessibilité au solvant) pour la validation expérimentale. Par ailleurs, nous avons étudié la propension des allergènes Der f 5 et Der f 21 de *Dermatophagoides farinae*, Der p 21 de *Dermatophagoides pteronyssinus*, Sui 5 m de *Suidasia medanensis* et Lep d 5 de *Lepidoglyphus destructor* et Blo t 21 de *Blomia tropicalis* à dimériser et former une grande cavité hydrophobe similaire à celle de Der p 5. Cette cavité hydrophobe a été suggérée dans la littérature comme pouvant intervenir dans l'initiation du processus allergique en liant certains ligands hydrophobes. Nos résultats suggèrent que seulement Der f 5 pourrait former un dimère semblable à Der p 5. Les autres allergènes (Blo t 5, Sui 5 m, Lep d 5, Der p 21, 21 et Der f Blo t 21) ne présentent pas une telle propension. En outre, nous avons identifié des mutations qui devraient déstabiliser et/ou entraver la formation de la structure dimérique de Derp5, ce qui pourrait aider à la compréhension du rôle du processus de dimérisation dans la réponse allergique de ces allergènes.

Mots clés: allergies aux acariens, allergènes des familles 5 et 21, épitopes, dimérisation.

## Abstract:

In this research we focused on the *in silico* study (using bioinformatics approaches and molecular modelling) of the 3D structures of house dust mite allergens from families 5 and 21. We aimed to define the epitopes of Der p 5 and provided information about their dimerization propensity. This contributes to the study of their mechanisms of allergenicity and may lead to the design of hypoallergenic variants for immunotherapy. Four antigenic regions were identified for Der p 5: H34-L39 (epitope-Nt), H58-D69 (epitope II), L92-E107 (epitope I) and K129-V132 (epitope-Ct) (according to the PDB 3MQ1 numbering of Der p 5). Subsequently, residues located in these regions that might play a crucial role in binding to immunoglobulins E (IgE) were selected according to criteria that we defined (nature of amino acids and solvent accessibility) for experimental validation. Further, we have studied the propensity of the allergens Der f 5 and Der f 21 from *Dermatophagoides farinae*, Der p 21 from *Dermatophagoides pteronyssinus*, Sui m 5 from *Suidasia medanensis*, Lep d 5 from *Lepidoglyphus destructor* and Blo t 21 from *Blomia tropicalis* to dimerize and form a large cavity in the same way as Der p 5. This hydrophobic cavity has been suggested in the literature to be involved in the allergic process through the binding to hydrophobic ligands. Our results suggest that only Der f 5 could form a dimer similar to Der p 5. The other allergens (Blo t 5, Sui m 5, Lep d 5, Der p 21, Der f 21 and Blo t 21) did not present such a propensity. Moreover, we identified mutations that should destabilize and/or prevent the formation of the Derp5 dimeric structure, which may help us to understand the role of the dimerization process in the allergic response of these allergens.

Keywords: dust mite allergy, allergens from families 5 and 21, epitopes, dimerization.

## ملخص:

نسعى في هذا البحث إلى الدراسة في السيليكو (*in silico*) (باستعمال المناهج المعلوماتية والنمذجة الجزيئية) للهيكل الثلاثة الأبعاد لمسببات الحساسية لعث الغبار من الأسر 5 و 21. نهدف إلى تعريف حواتم مسبب الحساسية Der p 5 وتوفير المعلومات حول سلوك مسببات الحساسية من الأسر 5 و 21 من خلال دراسة عملية "Dimérisation" (تشكيل الثنائيات الجزيئية). وهذا يساهم في دراسة آلياتها من الحساسية، كما قد يؤدي إلى تصميم جزيئات متغيرات من هذه المسببات الحساسية للعلاج المناعي. حددت 4 حواتم لـ Der p 5: H39-L34 (حاتم Nt)، H58-D69 (حاتم II)، L92-E107 (حاتم I)، K129-V132 (حاتم Ct) (وفقا لترقيم الأحماض الأمنية الموجودة في ملف هيكل "3MQ1"). فيما بعد، تم إنتقاء وفقا لمعاييرنا (طبيعة الأحماض الأمنية ووصول المادة المذيبة إليها)، الأحماض الأمنية التي يمكن أن تغلب دورا حاسما في هذه الحواتم في الارتباط مع مضادات الحساسية من نوع "IgE". من ناحية أخرى، قمنا بدراسة ميل مسببات الحساسية Der f 5 و Der f 21 و Der p 21، Sui m 5 و Lep d 5 المعزولين من عدة عث لتشكيل ثنائيات جزيئية وتكوين تجويف كاره للماء بنفس الطريقة مثل Der p 5. تشير نتائجنا أن Der f 5 وحده باستطاعته أن يشكل مركبا ثنائي الجزيئية مماثلا لـ Der p 5، أما مسببات الحساسية الأخرى (Der p 21، Der f 21، Sui m 5 و Lep d 5، Blo t 5 و Blo t 21) فلم تظهر هذا الميل. إضافة إلى هذا، قمنا بتحديد الطفرات التي يمكن أن تفقد الاستقرار أو تمنع تشكيل الجزيئات الثنائية الهيكل، التي قد تساعدنا في فهم دور عملية تشكيل الثنائيات في الإستجابة لهذه المواد المثيرة للحساسية. كلمات البحث : الحساسية لعث الغبار، مسببات الحساسية من الأسر 5 و 21، حواتم، تشكيل الثنائيات الجزيئية.

# Sommaire

	Page
<b>Introduction générale</b>	1
<b>Chapitre I: Synthèse bibliographique</b>	
I.1 Allergie : Définitions et nomenclatures .....	6
I.2 Allergie aux acariens .....	7
I.2.1 Taxonomie et écologie des acariens .....	7
I.2.2 Allergènes des acariens .....	9
I.2.3 Mécanisme d'activation du système immunitaire par les allergènes d'acariens .....	12
I.2.4 Rôle de la structure des allergènes dans leur allergénicité .....	19
I.3 Traitement des allergies respiratoires par l'immunothérapie spécifique .....	27
I.4 Méthodes Bioinformatiques d'aide à la compréhension des mécanismes liés à l'allergie .....	29
I.4.1 Bases de données généralistes et spécialisées .....	30
I.4.2 Méthodes de prédiction des épitopes B .....	33
I.4.3 Prédiction de la structure des allergènes.....	37
I.4.4 Amarrage protéine-protéine/protéine-ligands.....	38
<b>Chapitre II: Etude de la réactivité croisée entre les allergènes des familles 5 et 21 et Prédiction des épitopes de Der p 5</b>	
II.1 Introduction et problématique .....	44
II.2 Méthodes .....	49
II.2.1 Etude de la réactivité croisée entre les allergènes des familles 5 et 21.....	49
II.2.1.1 Alignement de séquences .....	49
II.2.2 Prédiction des épitopes de Der p 5 .....	49
II.2.2.1 Structures expérimentales des allergènes Der p 5, Blo t 5 et Blo t 21.....	50
II.2.3 Criblage des résidus épitopes majeurs de Der p 5 pour la validation expérimentale.	51
II.2.4 Visualisation des structures.....	52
II.3 Résultats et discussions .....	53
II.3.1 Analyse de l'alignement de séquences.....	53
II.3.2 Prédiction des épitopes de Der p 5.....	60
II.3.2.1 Criblage des résidus épitopes majeurs de Der p 5 pour la validation expérimentale.....	61
II.3.2.2 Comparaison des épitopes prédits de Der p 5 avec leurs correspondants dans la séquence de Der f 5.....	66
II.4 Conclusions.....	67

<b>Chapitre III: Modélisation des structures tridimensionnelles monomériques et oligomériques des allergènes d'acariens des familles 5 et 21</b>	
III.1 Introduction et problématique .....	69
III.2 Méthodes .....	73
III.2.1 Modélisation comparative des structures 3D des allergènes des familles 5 et 21....	73
III.2.1.1 Création des modèles .....	73
III.2.1.2 Evaluation de la qualité et sélection des modèles 3D .....	74
III.2.2 Homodimérisation .....	75
III.2.2.1 Simulations d'amarrage protéine-protéine .....	75
III.2.2.2 Calcul de l'écart quadratique moyen (RMSD) .....	75
III.2.2.3 Assignation des interactions à l'interface des dimères.....	76
III.2.2.4 Identification des cavités.....	76
III.3 Résultats et discussions.....	77
III.3.1 Analyse des résidus de l'interface de Der p 5 dans l'alignement multiple de séquence.....	77
III.3.2 Modélisation comparative de la structure des monomères.....	80
III.3.3 Homodimérisation.....	83
III.3.3.1 Fiabilité de ClusPro pour prédire la structure des dimères .....	83
III.3.3.2 Prédiction de la structure dimérique des allergènes et analyse de leurs propriétés.....	87
III. 4 Conclusions.....	95
<b>Conclusions générales et perspectives</b>	97
<b>Références bibliographiques</b>	102
<b>Annexes</b>	140

## Liste des abréviations

**3D:** Tridimensionnel

**CAPRI:** *Critical Assessment of PRediction of Interactions*

**CASP:** *Critical Assessment of techniques for protein Structure Prediction*

**CD:** Cluster de Différenciation.

**CXCL:** *Chemokine (C-X-C motif) ligand*

**DC-SIGN:** *Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin*  
(Molécules d'adhésion intracellulaire non-intégrine 3 spécifiques aux cellules dendritiques)

**DSSP :** *Dictionary of Secondary Structure of Proteins*

**FFT:** *Fast Fourier Transform* (Transformée de Fourier rapide)

**GM-CSF:** *Granulocyte macrophage colony-stimulating factor* (Facteur stimulant les colonies de granulocytes et de macrophages).

**ICAM-3:** *Intercellular adhesion molecule-3* (Molécule d'adhésion intercellulaire-3)

**IgE:** Immunoglobuline E

**IL:** Interleukine

**INT- $\gamma$ :** Interféron gamma

**LPS:** Lipopolysaccharides

**MD-2:** *Myeloid Differentiation 2*

**PAR-2:** *Proteinase Activated Receptors 2* (Récepteur activés par les protéases 2)

**PDB:** *Protein Data Bank*

**RMN:** Résonance Magnétique Nucléaire

**RMSD:** *Root Mean Square Deviation*

**SP:** Surfactant pulmonaire

**T<sub>H</sub>:** Lymphocytes T helper

**TLR:** Toll-like receptor

**TNF- $\alpha$ :** *Tumor necrosis factor-alpha* (Facteur-alpha de nécrose tumorale)

**UISI:** Union Internationale des Sociétés d'Immunologie

**SVM:** *Support Vector Machine* (Machine à vecteurs supports).

## Liste des figures

	Page
Figure n°1: Voie classique de la réponse allergique.....	13
Figure n°2: Molécules intervenant dans l'activation du système immunitaire en association avec les acariens.....	15
Figure n°3: Voies d'activation du système immunitaire par les allergènes d'acariens et leurs adjuvants.....	16
Figure n°4: Mécanisme d'action de l'allergène majeur Der p 1.....	17
Figure n°5: Structures tridimensionnelles des allergènes d'acariens du groupe 1.....	21
Figure n°6: Structures tridimensionnelles des allergènes d'acariens du groupe 2.....	22
Figure n°7: Structure tridimensionnelle cristallographique de Der p 7 (code PDB 3H4Z).	23
Figure n°8: Structure RMN de Der f 13 (code PDB 2A02) et position de ses épitopes.....	24
Figure n°9: Structures 3D des allergènes d'acariens des groupes 5 et 21.....	25
Figure n°10: Approches pour la génération de vaccins hypoallergéniques contre les allergies.....	29
Figure n°11: Mesures pour l'évaluation de l'amarrage protéine "A" - protéine "B" .....	40
Figure n°12: Localisation des épitopes de Blo t 5 et de Blo t 21 déterminés expérimentalement.....	47
Figure n°13: Alignement multiple de séquences des allergènes des familles 5 et 21 étudiés	55
Figure n°14: Distribution des résidus chargés dans les structures de Blo t 21 (code PDB 2LM9), de Blo t 5 (code PDB 2JMH) et de Der p 5 (code PDB 3MQ1_A)...	57
Figure n°15: Superposition de la structure au niveau des épitopes de Blo t 5, de Blo t 21 et de leurs correspondant dans Der p 5.....	58
Figure n°16: Alignement de séquence multiple entre Blo t 5, Der p 5 et Blo t 21 et positions de leurs épitopes B prédits par le programme SEPPA.....	60
Figure n°17: Position des épitopes prédits de Der p 5 par le programme SEPPA et résidus majeurs proposés, d'après nos critères de sélection, pour la validation expérimentale.....	64
Figure n°18: Représentation de la structure 3D des allergènes Blo t 5 (code PDB 2JMH), de Blo t 21 (code PDB 2LM9) et de Der p 5 (code PDB 3MQ1).....	70
Figure n°19: Alignement multiple des séquences des allergènes étudiées obtenu par ClustalW .....	78
Figure n°20: Superposition entre la structure dimérique expérimentale de Der p 5 '3MQ1_AB'(chaines A et B du code pdb 3MQ1) et la structure du dimère de Der p 5 prédit par ClusPro 'derp5_dimer[3MQ1_A]'.....	84
Figure n°21: Interactions de type (a) contacts hydrophobes, (b) liaisons hydrogène et (c) ponts salins calculées par PIC entre les résidus de l'interface des dimères 3MQ1_AB, derp5_dimer [3MQ1_A] et derf5_dimer5.....	85
Figure n°22: Structures du dimère de Der p 5 (3MQ1_AB) et des modèles de dimères prédites par ClusPro des allergènes des familles 5 et 21 étudiés.....	90
Figure n°23: Représentation de l'hélice N <sub>terminale</sub> coudée après superposition de la structure expérimentale de Der p 5 (3MQ1_A) avec les structures prédites de Blo t 5: le blot5_monomer <sub>6</sub> [3MQ1_A] et blot5_monomer <sub>2</sub> [3MQ1_A].....	93

## Liste des tableaux

		Page
Tableau n°1:	Principales espèces des acariens domestiques .....	8
Tableau n°2:	Rôles biologiques des allergènes d'acariens et leur fréquence de fixation d'IgE.....	11
Tableau n°3:	Bases de données généralistes.....	30
Tableau n°4:	Bases de données spécialisées en allergies.....	32
Tableau n°5:	Outils de prédictions des épitopes B.....	35
Tableau n°6:	Liste non-exhaustive des programmes d'amarrage protéine-protéine/protéine-ligands les plus populaires.....	42
Tableau n°7:	Réactivité croisée entre les allergènes des familles 5 et 21.....	45
Tableau n°8:	Pourcentage d'identité et de similarité de séquences entre les allergènes étudiés.....	53
Tableau n°9:	Valeurs de la moyenne de l'accessibilité au solvant +/- écart-type (<%acc.> +/- $\sigma$ ) des résidus épitopes sauvages de Blo t 5 et de Blo t 21 déterminés expérimentalement, et de l'accessibilité au solvant (acc.) des résidus épitopes des deux chaînes A et B de Der p 5 prédits par SEPPA. La valeur moyenne de la différence d'énergie libre (< $\Delta\Delta G$ >) +/- écart-type (( $\Delta\Delta G$ ) +/- $\sigma$ ) pour Blo t 5 et Blo t 21 et la valeur de la différence d'énergie libre pour Der p 5 ( $\Delta\Delta G$ ) après leur mutation en alanine sont également fournies.....	63
Tableau n°10:	Pourcentages d'identité de séquences entre Der p 5 et les autres allergènes calculés au niveau des résidus impliqués dans les interactions à l'interface du dimère 3MQ1_AB.....	79
Tableau n°11:	Valeurs d'énergie, calculées avec Anolea et Gromos, des modèles monomériques générés sur la base des patrons 2JMH, 2LM9 et 3MQ1_A..	82
Tableau n°12:	Pourcentage des interactions communes à l'interface de 3MQ1_AB, de derp5_dimer[3MQ1_A] et de derf5_dimer.....	86
Tableau n°13:	Valeurs de la fonction de score de ClusPro des meilleurs dimères prédits pour les allergènes étudiés en utilisant 3MQ1_A comme patron et valeurs du rmsd obtenues après superpositions de ces modèles avec la structure cristallographique 3MQ1_AB.....	89

# *Introduction générale*

Les maladies allergiques constituent un problème majeur de santé publique. L'Organisation Mondiale de la Santé (OMS) les classe au quatrième rang mondial des pathologies. Les allergies respiratoires, essentiellement la rhinite allergique, l'asthme et les dermatites allergiques sont les manifestations les plus fréquentes. D'après les estimations de l'OMS, il y a actuellement 235 millions d'asthmatiques dans le monde et des millions d'individus souffrent de rhinite allergique (OMS, 2012). La prévalence mondiale de ces allergies a été évaluée grâce à deux grandes enquêtes épidémiologiques, une chez l'adulte (étude ECRHS "*European Community Respiratory Health Survey*", ECRHS-I, 1996; ECRHS-II, 2002) et l'autre chez l'enfant (étude ISAAC "*International Study of Asthma and Allergies in Childhood*", ISAAC, 1998a; ISAAC, 1998b).

La prévalence de l'asthme et de la rhinite allergique chez les adultes âgés de 20 à 44 ans dans les 22 pays concernés par la première étude ECRHS (1991-1993) variait de 1,3% à 9,7% et de 9,5 à 40,9 %, respectivement (ECRHS-I, 1996; Bousquet *et al.*, 2008a). Les prévalences étaient particulièrement élevées dans les pays anglophones (la Nouvelle-Zélande, l'Australie, les Etats-Unis, l'Angleterre) et plus basses dans les pays méditerranéens (la Grèce, l'Italie, l'Espagne, le Portugal, l'Algérie) et en Europe de l'Ouest (la France, la Belgique, l'Allemagne, la Suisse et l'Autriche). L'étude ISAAC a été réalisée afin de comparer la prévalence de ces maladies respiratoires dans 56 pays chez les enfants de 6 à 7 ans et dans 38 pays chez les jeunes de 13 à 14 ans. Les variations de la prévalence étaient de 2,1 à 32,2% pour l'asthme, de 1,4% à 39,7% pour la rhinite allergique et de 0,3% à 20,5% pour la dermatite atopique (l'eczéma) (Strachan *et al.*, 1997; ISAAC, 1998a; ISAAC, 1998b). Une deuxième phase pour les deux études (ECRHS II et ISAAC II) a été réalisée pour identifier les facteurs déterminant les différences de prévalence observées (ECRHS-II, 2002; Weiland *et al.*, 2004; Bedada *et al.*, 2007).

Récemment, une troisième étude ISAAC a été effectuée. Les résultats de cette étude, comparés à ceux de la première phase, indiquent une diminution de la prévalence de l'asthme, en particulier chez les enfants âgés de 13-14 ans dans les pays anglophones et en Europe de l'Ouest, une légère augmentation de la prévalence en Amérique latine, Chine, Indonésie, Taïwan et Corée du Sud, et une très forte augmentation au Maroc, en Tunisie et en Algérie (Pearce *et al.*, 2007; Björkstén *et al.*, 2008).

La première enquête réalisée en Algérie (en 1991) dans la région d'Alger a permis d'évaluer la prévalence de l'asthme à 1,3% (Belhocine & Ait-Khaled 1991). L'étude ECRHS réalisée de

1991 à 1995 a retrouvé une prévalence de l'asthme de 3% chez les adultes et de 3,6% chez les adolescents de 13-14 ans (ECRHS-I, 1996). La prévalence chez ces derniers a doublé dans cette même population, atteignant 6,4% d'après l'étude ISAAC (ISAAC, 1998a; Bouayad et Afif 1998). La prévalence de la rhinite allergique était de 9,5% (ECRHS-I, 1996). En 2010, une étude de l'environnement allergique d'une population d'enfants asthmatiques a été réalisée à Annaba (située au Nord-Est de l'Algérie) (Boumendjel et *al.*, 2010). Cette étude a montré que, sur les 75 patients allergiques testés, la prédisposition génétique basée sur les critères familiaux est très fréquente (74%), et que 96% de ces enfants allergiques sont asthmatiques dont 46% ont un asthme sévère (Boumendjel et *al.*, 2010). Aujourd'hui, l'Algérie compte plus d'un million d'asthmatiques et près de trois millions de malades souffrant de rhinites allergiques (Douagui, 2007).

Ces maladies respiratoires sont déclenchées lors de l'exposition à des antigènes ou allergènes. Ces antigènes sont:

- \* soit des pneumallergènes aéroportés ou inhalés comme les substances des acariens, les pollens, les moisissures et les poils d'animaux domestiques,
- \* ou des trophallergènes d'origine alimentaire issus d'arachides ou du lait par exemple,
- \* ou des allergènes chimiques tels le formol, le latex ou les antibiotiques.

Une étude basée sur l'enquête ECRHS a montré que les sensibilités à l'acarien *Dermatophagoides pteronyssinus*, aux pollens et les poils de chat sont les plus prédominantes (Bousquet et *al.*, 2007).

Plusieurs autres études ont montré que l'allergie aux acariens de poussière de maison, essentiellement due à *Dermatophagoides pteronyssinus*, *Dermatophagoides farinae*, *Euroglybphus maynei* et *Blomia tropicalis*, est la cause la plus répandue de la rhinite allergique, de l'asthme et de la dermatite atopique (Arruda et *al.*, 1997; Tsai et *al.*, 2003; Simpson et *al.*, 2003). Elle est responsable de 50% des manifestations allergiques dans le monde (Jacquet, 2009). Cette allergie aux acariens est très fréquente en Algérie, surtout dans les villes côtières où l'humidité très élevée favorise la prolifération des acariens (Boumendjel et *al.*, 2010).

A l'heure actuelle, divers allergènes d'acariens ont été purifiés. Ces allergènes sont groupés en familles selon leur fonction biologique et l'ordre chronologique de leur purification. Un allergène est caractérisé par ses épitopes (déterminants antigéniques). Il est dit majeur s'il

déclenche une réaction chez au moins 50% des patients allergiques testés. Certains de ces allergènes ont une activité sérine protéase, d'autres sont des trypsines, des alpha-amylases, des chitinases ou des fixateurs des acides gras. Cette fonction joue un rôle très important dans la compréhension du mécanisme de déclenchement de la réponse immunitaire contre ces allergènes.

Plusieurs allergènes appartenant à la famille 5 ont été isolés à partir de différentes espèces d'acariens: Der p 5 de *Dermatophagoides pteronyssinus*, Blo t 5 de *Blomia tropicalis*, Der f 5 de *Dermatophagoides farinae*, Lep d 5 de *Lepidoglyphus destructor*, Ale o 5 de *Aleuroglyphus ovatus*, Tyr p 5 de *Tyrophagus putrescentiae*, Sui m 5 de *Suidasia medanensis* et Gly d 5 de *Glycyphagus domesticus*. Ces allergènes présentent des similarités de séquence avec les trois allergènes de la famille 21 connus jusqu'à présent (Der p 21, Blo t 21 et Der f 21). En outre, la fonction biologique de ces allergènes est, à l'heure actuelle, inconnue, et peu d'informations sur la caractérisation de leurs épitopes et la réactivité croisée entre ces allergènes sont disponibles.

Une voie permettant la compréhension de la fonction biologique et des propriétés d'un allergène passe par l'étude de sa structure tridimensionnelle (3D). Parmi les allergènes des familles 5 et 21, seules les structures de Blo t 5, de Blo t 21 et de Der p 5 ont été déterminées expérimentalement par résonance magnétique nucléaire (RMN) ou par cristallographie aux rayons-X (RX) (Naik et al., 2008; Chan et al., 2008; Mueller et al., 2010a; Tan et al., 2012). Leurs structures sont composées de trois hélices alpha formant une organisation en *coiled-coil*. Les structures de Blo t 5 et de Blo t 21 sont monomériques tandis que Der p 5 a été trouvée comme étant un homodimère formant une cavité hydrophobe au centre de la molécule.

D'autres allergènes d'acariens ont des cavités hydrophobes dans leurs structures 3D. Ces cavités leur permettent de fixer des ligands hydrophobes, essentiellement des acides gras ou des lipopolysaccharides (LPS), présents dans la membrane externe des bactéries Gram négatives et d'autres molécules microbiennes. Les exemples les plus étudiés sont les allergènes d'acarien de la famille 2 (Der p 2 et Der f 2) (Derwenda et al., 2002 ; Suzuki et al., 2006). Ces allergènes présentent une similarité structurale avec les protéines Niemann-Pick type 2 qui transportent le cholestérol (Friedland et al., 2003) et avec la protéine MD2 (*Myeloid Differentiation 2*), qui joue un rôle dans la reconnaissance des LPS et dans leur présentation aux récepteurs TLR4 (*Toll-Like receptor 4*), ce qui déclenche la réponse

immunitaire (Park *et al.*, 2009). La RMN et les dosages biochimiques indiquent que Der p 7 (l'allergène de famille 7 de *Dermatophagoides pteronyssinus*) présente une cavité hydrophobe qui ne lie pas les LPS, mais lie avec une faible affinité la polymyxine B, un lipopeptide d'origine bactérienne (*Bacillus polymixa*) (Mueller *et al.*, 2010b).

De nombreuses recherches cliniques ont été menées afin d'étudier la relation entre la concentration en endotoxines et les  $\beta$ -D-glucanes dans l'environnement (particulièrement les poussières de maisons) et la prévalence et la sévérité de l'asthme (Thorn et Rylander, 1998; Douwes *et al.*, 2000 ; Park *et al.*, 2001 ; Smit *et al.*, 2008). Il a été démontré que l'exposition à ces biomolécules exacerbe les symptômes de l'asthme essentiellement chez les patients allergiques aux acariens (Ferández-Caldas, 2002; Rizzo *et al.*, 2007). Des années plus tard, Valerio *et al.* (2005) ont pu séquencer, à partir d'extraits d'acariens, des gènes d'ARN 16S de certaines bactéries Gram négatives (certaines espèces de *Bartonella*, *Escherichia coli*, certaines espèces de *Pseudomonas* et de *Acinetobacter*) qui se trouvent être la source des endotoxines trouvés en complexe avec les allergènes des acariens et aggravent l'allergie.

Notre thèse s'inscrit dans le cadre de l'étude *in silico* des structures 3D des allergènes d'acariens de poussières de maison des familles 5 et 21. Nous avons pour objectifs de (1) définir les épitopes de Der p 5 et (2) de fournir des informations quant au comportement moléculaire de ces allergènes en étudiant leur phénomène de dimérisation afin de contribuer à l'étude des mécanismes de leur allergénicité.

Ce travail est organisé en trois chapitres. Le premier chapitre sera consacré à une revue de littérature sur les allergies respiratoires, les allergènes des acariens et leurs fonctions biologiques. En outre, nous aborderons également les méthodes bioinformatiques d'aide à la compréhension des mécanismes liés à l'allergie, les méthodes de prédiction des épitopes et celles de prédiction de structures et de fonctions des protéines.

Dans le deuxième chapitre, nous exposerons les méthodes et les résultats de la prédiction des épitopes B de l'allergène Der p 5 dans le but d'identifier les résidus impliqués dans la liaison avec les immunoglobulines E (IgE) qui conditionne l'allergénicité de cet allergène et de discuter la réactivité croisée entre les allergènes des familles 5 et 21.

Le troisième chapitre, sera dédié à l'étude bioinformatique de la structure des allergènes de familles 5 et 21 et de leur état oligomérique. Des techniques *in silico* de prédiction de structures 3D de protéine et d'amarrage protéine-protéine seront utilisées afin d'étudier la

structure 3D des allergènes des familles 5 et 21, ainsi que leur propension à former des oligomères similaires à ceux observés chez Der p 5. En effet, le dimère de cet allergène présente une cavité hydrophobe susceptible de fixer des ligands. Une hypothèse avancée et non encore vérifiée est que cette cavité pourrait être impliquée dans la réponse immunitaire, suite à l'éventuelle liaison à un ligand (Mueller et *al.*, 2010a; Thomas, 2012).

Enfin, nous terminerons cette thèse par une conclusion générale et des perspectives.

*Chapitre I:*  
*Synthèse bibliographique*

Selon l'Organisation Mondiale de la Santé, les allergies se situent au quatrième rang des maladies dans le monde. Plus d'un enfant sur cinq est allergique.

### **I.1 Allergie : Définitions et nomenclatures**

L'allergie ou l'hypersensibilité est une réaction excessive et pathologique du système immunitaire, déclenchée suite à un contact avec une substance généralement étrangère à l'organisme appelée antigène ou allergène (Johansson et *al.*, 2001).

L'hypersensibilité immédiate, dite de type 1 selon la classification de Gell et Coombs (1963), est la forme la plus fréquente de l'allergie. Elle affecte plus de 25% de la population dans les pays industrialisés (Larché et *al.*, 2006). Cette hypersensibilité immédiate est définie comme étant l'ensemble des phénomènes résultant de l'interaction d'un antigène avec des anticorps (les immunoglobulines E - IgE) fixés sur les mastocytes, granulocytes basophiles et macrophages. Cela provoque la libération de médiateurs chimiques comme l'histamine et les cytokines (interleukines) responsables des processus aussi variés que les allergies respiratoires (asthme et rhinite allergiques), le choc anaphylactique, diverses allergies oculaires et la dermatite atopique (eczéma) (Galli et *al.*, 2008; Bidad et *al.*, 2011).

Le terme atopie introduit pour la première fois par Coca & Cooke (1923) est par définition une prédisposition génétique (héréditaire) à développer certaines maladies allergiques comme l'asthme allergique, la rhinite allergique (rhume des foins) et l'eczéma (Caruso et *al.*, 2009). L'atopie est caractérisée par des élévations des taux d'IgE sériques en réponse à des substances communes de l'environnement (tels que les acariens, les pollens de graminées et les allergènes alimentaires) (Johansson et *al.*, 2001).

L'asthme, quant à lui, est une affection inflammatoire chronique des voies respiratoires inférieures caractérisée par l'obstruction réversible des voies aériennes. Chez les personnes sensibles, l'inflammation provoque des crises récurrentes de respiration sifflante, de toux, d'essoufflement, et d'oppression thoracique (Bousquet et *al.*, 2000). Par ailleurs, la rhinite allergique est un trouble des voies aériennes supérieures (au-dessus du larynx) résultant de l'inflammation médiée par les immunoglobulines E (IgE) du nez lors d'un contact de la muqueuse nasale avec les allergènes. Ces symptômes comprennent la rhinorrhée, le prurit nasal, l'éternuement et l'obstruction nasale (Bousquet et *al.*, 2000).

L'asthme et la rhinite allergique se produisent fréquemment de façon concomitante. En effet, la présence de la rhinite allergique précède souvent le développement de l'asthme et est un facteur de risque connu pour l'asthme (Thomas, 2006; Bousquet et *al.*, 2008b; Bousquet et *al.*, 2012).

Ces réactions allergiques sont déclenchées par les allergènes trouvés dans l'environnement comme le pollen ou les poussières des maisons composées de squames animales, poils de chat et de chiens et surtout d'acariens. L'allergie aux acariens de poussières de maisons (*House dust mite, HDM*) est la cause la plus répandue de l'asthme, de la rhinite (Thomas et *al.*, 2010) et de la dermatite atopique (Shek et *al.*, 2010; Fuiano & Incorvaia, 2012). Plus de 85% d'asthmatiques dans le monde sont allergiques aux acariens de poussières de maisons (Gregory & Lloyd, 2011).

## **I.2 Allergie aux acariens**

La poussière de maison représente la première source d'allergènes de l'environnement domestique. Plusieurs études ont démontré l'existence d'une relation nette entre l'état clinique de l'allergie et les quantités d'acariens présentes dans son environnement. En effet, l'exposition à des quantités supérieures à 10 microgrammes/g-de-poussière d'allergènes majeurs des acariens représente un facteur de risque d'exacerbation des symptômes d'asthme et d'allergie respiratoire chez les sujets sensibilisés (Surdu et *al.*, 2006).

### **I.2.1 Taxonomie et écologie des acariens**

Les acariens appartiennent à l'embranchement des arthropodes (*Arthropoda*), la classe des arachnides (*Arachnida*), la sous-classe des Acariens (*Acari*) et l'ordre des Astigmatés (*Astigmata*) (Spieksma, 1997) qui comprend quatre familles ayant une importance capitale dans l'origine des phénomènes allergiques chez l'homme: *Pyroglyphidae*, *Acaridae*, *Glyciphagidae* et *Echimyopodidae*. Les principales espèces de chaque famille sont représentées dans le tableau n°1.

Tableau n°1: Principales espèces des acariens domestiques

Familles	Espèces	Source
<i>Pyroglyphidae</i>	<i>Dermatophagoides pteronyssinus</i> <i>Dermatophagoides farinae</i> <i>Dermatophagoides siboney</i> <i>Euroglyphus maynei</i>	Poussières de maisons
<i>Acaridae</i>	<i>Acarus siro</i> <i>Tyrophagus putrescentiae</i> <i>Aleuroglyphus ovatus</i> <i>Suidasia medanensis</i>	Lieux de stockage
<i>Glyciphagidae</i>	<i>Glycophagus domesticus</i> <i>Lepidoglyphus destructor</i> <i>Lepidoglyphus domesticus</i>	Lieux de stockage
<i>Echimyopodidae</i>	<i>Blomia tropicalis</i> <i>Blomia kulagini</i>	Lieux de stockage /poussières de maisons Lieux de stockage

Les acariens sont ubiquitaires. Ils ont une taille microscopique variant de 170 à 500 µm et selon leur source, on distingue deux catégories: les acariens de poussières de maison trouvés surtout dans la literie, en particulier le matelas, les tapis, les moquettes et les meubles en se nourrissant principalement de débris organiques et de squames cutanées (Zock et al., 1994 ; Arbes et al., 2003) et les acariens de stockage trouvés essentiellement dans les entrepôts de stockage des denrées alimentaires (grains, graminées, farine, paille, foin) (Arlian, 2002). Les acariens de la poussière des maisons appartiennent à la famille des *Pyroglyphidae*, tandis que les acariens de stockage appartiennent aux trois familles *Acaridae*, *Glyciphagidae* et *Echimyopodidae* (Yadav et al., 2006). L'acarien de stockage *Blomia tropicalis*, autrefois classé parmi les *Glyciphagidae*, a été récemment intégré dans la famille des *Echimyopodidae* et considéré comme un acarien domestique des poussières de maisons des régions tropicales et subtropicales (Wong et al., 2011).

La distribution de ces acariens domestiques varie selon le climat, la zone géographique et les conditions socio-économiques (Tang et al., 2011). La température et l'humidité sont des facteurs cruciaux pour leur croissance (Lintner & Brame, 1993). Ils se développent dans les environnements chauds et humides avec un optimum de température de 20–25°C et une humidité relative comprise entre 65 et 75% (Arlian et al., 2001). Dans les climats tempérés on trouve essentiellement *Dermatophygoïdes pteronyssinus* et *Dermatophagoïdes farinae*, dans les régions humides comme l'Angleterre ou le Sud des Etats-Unis on trouve *Euroglyphus maynei*, et dans les régions tropicales et subtropicales c'est surtout *Blomia tropicalis* qui

prédomine (Morgan & Arlian, 1997; Arlian, 2002).

En Algérie, peu d'études approfondies ont été réalisées sur l'allergie aux acariens, même si celle-ci est très fréquente. Une étude préliminaire de la prévalence des acariens de poussières de maisons a été réalisée dans 4 régions d'Algérie (Alger, Oran, Béjaïa et Tizi-Ouzou). Les résultats ont montré l'existence d'une acarofaune variée où *Dermatophagoides pteronyssinus* est l'acarien dominant et permanent suivi de *Dermatophagoides farinae* et d'*Euroglyphus maynei*. D'autres acariens appartenant à d'autres familles ont été également identifiés (*Glycyphagus sp.*, *Tyrophagus sp.*, *Cheyletus sp.* et *Chortoglyphus arcuatus*) (Abed-Benamara, 1983; Bouziane et al., 2001).

Les acariens incriminés dans les allergies respiratoires et les dermatites atopiques sont principalement *Dermatophagoides pteronyssinus*, *Dermatophagoides farinae* et *Blomia tropicalis*. Dans la plupart des études publiées sur la prévalence de sensibilisation aux aéroallergènes dans le monde, *Dermatophagoides pteronyssinus* est souvent l'acarien prédominant (Nelson et al., 1996; Carabello et al., 1998; Weghofer et al., 2008a; Shek et al., 2010; Kidon et al., 2011; Zhang et al., 2011; Soto-Quiros et al., 2012; Chruszcz et al., 2012). Les acariens de stockage sont reconnus pour être la cause des dermatites de contact professionnel (comme la dermatite des boulangers provoquée par *Acarus siro*) et à l'origine des allergies alimentaires sévères (Sánchez-Ramos et al., 2007). Yi et al. (2009) ont démontré que l'ingestion de la farine contaminée par les acariens peut déclencher une réaction anaphylactique sévère dans les régions tropicales et subtropicales. Le rôle de ces acariens de stockage en allergie respiratoire a été reconnu plus tardivement que celui des acariens de poussières de maisons (Kronqvist et al., 1999; Chambers et al., 1999).

### **I.2.2 Allergènes des acariens**

Les sources des allergènes des acariens sont les particules fécales qui s'accumulent dans les poussières, mais aussi les sécrétions salivaires et les débris d'acariens morts (Nadchatram, 2005). Il faut savoir que chaque acarien a ses propres allergènes. Leur particularité est qu'il existe une réactivité croisée entre les allergènes issus de différents acariens (Arlian et al., 2009). En effet, les anticorps de type IgE se fixent sur une partie spécifique de la séquence appelée déterminant antigénique ou épitope. Lorsqu'un même anticorps est susceptible de se lier à des allergènes différents mais dont les épitopes sont proches ou identiques on parle de réactivité croisée entre ces allergènes.

Les allergènes d'acariens sont soit des protéines ou des glycoprotéines avec un poids moléculaire allant de 11 à 190 kDa. Leur nomenclature systématique est dictée par l'Union Internationale des Sociétés d'Immunologie (UISI) et contient trois parties séparées chacune par un espace: les trois premières lettres du genre de l'organisme d'où l'allergène est isolé, suivi de la première lettre de l'espèce, et enfin un chiffre arabe correspondant à l'ordre chronologique d'identification de l'allergène et de sa fonction biologique (Chapman et *al.*, 2007). Donc, **Der p 1** est le **premier** allergène isolé à partir de l'acarien *Dermatophagoides pteronyssinus*. Les allergènes natifs (naturels) ont le préfixe "n" devant leurs noms (ex., nDer p 1), alors que les allergènes recombinants sont désignés par un "r" (ex., rDer p 1).

D'autres types moléculaires d'allergènes peuvent également être rencontrés: les iso-allergènes et les isoformes. Par définition, un iso-allergène est un allergène de même masse moléculaire et fonction biologique qu'un autre mais avec une identité de séquence d'au moins 67% (Elsayed & Aas, 1970; Chapman, 2008), alors que le terme "isoformes" est employé pour désigner des variants polymorphiques du même allergène ayant plus de 90 % d'identité de séquences (Chapman, 2008). Si des iso-allergènes sont identifiés, un suffixe de 01 à 99 est utilisé (ex., Der p 1. 01, etc...). Par ailleurs, pour les variants moléculaires ou les isoformes, un chiffre indiquant le nombre des isoformes de ces allergènes est ajouté juste après le suffixe des iso-allergènes (par exemple, les 23 isoformes de Der p 1 sont désignés par Der p 1. 0101, Der p 1. 0102, ... Der p 1. 0123) (Piboonpocanun et *al.*, 2006; Chapman et *al.*, 2007). Les substitutions d'acides aminés dans ces séquences peuvent jouer un rôle très important pour le repérage des épitopes ou la compréhension des réactivités croisées entre les allergènes. En effet, les variants naturels qui présentent moins d'allergénicité reflètent les sites de fixations des IgE des allergènes sauvages. C'est pour cette raison que les noms des allergènes sont présentés de façon plus précise afin de distinguer telle ou telle isoforme (Chapman, 2008; Zakzuk et *al.*, 2009).

A l'heure actuelle, divers allergènes d'acariens ont été purifiés (tableau n°2). Certains allergènes ne sont pas connus mais ont été identifiés et leurs séquences sont répertoriées dans la banque de donnée des allergènes "WHO/IUIS" (<http://www.allergen.org/>) ou dans la banque de données de séquences nucléiques "Genbank" (<http://www.ncbi.nlm.nih.gov/genbank/>; Benson et *al.*, 2011). Par ailleurs, peu d'allergènes d'acariens ont une structure tridimensionnelle déterminée expérimentalement et déposée dans la banque de données *Protein Data Bank* "PDB" ([www.rcsb.org](http://www.rcsb.org); Berman et *al.*, 2000).

Tableau n°2: Rôles biologiques des allergènes d'acariens et taux de fixation d'IgE

Famille	Fonction biologique	Allergènes	Structure (pdb)	Taux de Fixation d' IgE (%)	Références
1	Cystéine protéase	Der p 1	2AS8	70-90	Chua <i>et al.</i> , 1988; de Halleux <i>et al.</i> , 2006
		Der f 1	3DS6		Dilworth <i>et al.</i> , 1991; Chruszcz <i>et al.</i> , 2009
		Blo t 1	-		Mora <i>et al.</i> , 2003; Ramos <i>et al.</i> , 2004
		Eur m 1	-		Smith <i>et al.</i> , 1999
2	Transporteurs de ligands	Der p 2	1KTJ	60-90	Chua <i>et al.</i> , 1990; Derewenda <i>et al.</i> , 2002
		Der f 2	2F08		Trudinger <i>et al.</i> , 1991; Suzuki <i>et al.</i> , 2006
		Blo t 2	-		WHO/IUIS/GenBank
		Eur m 2	-		Smith <i>et al.</i> , 1999
		Lep d 2	-		Eriksson <i>et al.</i> , 2001
		Tyr p 2	-		Eriksson <i>et al.</i> , 1998
3	Trypsine	Gly d 2	-	51-90	Gafvelin <i>et al.</i> , 2001
		Der p 3	-		Smith <i>et al.</i> , 1994
		Der f 3	-		Maruo <i>et al.</i> , 1997
		Blo t 3	-		Cheong <i>et al.</i> , 2003
		Eur m 3	-		Stewart <i>et al.</i> , 1992
4	Alpha amylase	Tyr p 3	-	25-46	Liao <i>et al.</i> , 2009
		Der p 4	-		Lake <i>et al.</i> , 1991; Mills <i>et al.</i> , 1999
		Blo t 4	-		Cheong <i>et al.</i> , 2009
5	Fonction inconnue	Eur m 4	-	9-70	Mills <i>et al.</i> , 1999
		Der p 5	3MQ1		Weghofer <i>et al.</i> 2008b; Mueller <i>et al.</i> , 20010a
		Der f 5	-		Cui <i>et al.</i> , 2012
		Blo t 5	2JMH		Caraballo <i>et al.</i> , 1996; Naik <i>et al.</i> , 2008
		Lep d 5	-		Eriksson <i>et al.</i> , 2001
		Ale o 5	-		WHO/IUIS/GenBank
		Sui m 5	-		WHO/IUIS/GenBank
		Gly d 5	-		WHO/IUIS/GenBank
6	Chymotrypsine	Tyr p 5	-	30-40	WHO/IUIS/GenBank
		Der p 6	-		Bennett & Thomas 1996
		Der f 6	-		Kawamoto <i>et al.</i> , 1999
7	Fonction inconnue	Blo t 6	-	50-62	WHO/IUIS/GenBank
		Der p 7	3H4Z		Shen <i>et al.</i> , 1993; Mueller <i>et al.</i> , 20010b
		Der f 7	-		Shen <i>et al.</i> , 2011
8	Glutathione S-transferase	Lep d 7	-	40	Eriksson <i>et al.</i> , 2001
		Der p 8	-		O'Neill <i>et al.</i> , 1995
9	Collagénase Serine protéase	Der p 9	-	>90	King <i>et al.</i> , 1996
10	Tropomyosine	Der p 10	-	5-80	Asturias <i>et al.</i> , 1998
		Der f 10	-		Aki <i>et al.</i> , 1995
		Blo t 10	-		Yi <i>et al.</i> , 2002
		Lep d 10	-		Saarne <i>et al.</i> , 2003
		Tyr p 10	-		Jeong <i>et al.</i> , 2007
11	Paramyosine	Der p 11	-	80	Lee <i>et al.</i> , 2004
		Der f 11	-		Tsai <i>et al.</i> , 1998
		Blo t 11	-		Ramos <i>et al.</i> , 2001
12	Fonction inconnue	Blo t 12	-	50	Puerta <i>et al.</i> , 1996
13	Protéines fixatrices d'acides gras	Der p 13	-	10-23	WHO/IUIS/GenBank
		Der f 13	2A0A		Chan <i>et al.</i> , 2006
		Blo t 13	-		U.S. WHO/IUIS/GenBank
		Lep d 13	-		Eriksson <i>et al.</i> , 2001
		Tyr p 13	-		Jeong <i>et al.</i> , 2005
14	Apolipoporphine	Der p 14	-	30-70	Epton <i>et al.</i> , 2001
		Der f 14	-		Aki <i>et al.</i> , 1994; Fujikawa <i>et al.</i> , 1996
		Eur m 14	-		WHO/IUIS/GenBank
15	Chitinase	Der p 15	-	70	O'Neil <i>et al.</i> , 2006
		Der f 15	-		McCall <i>et al.</i> , 2001
16	Gelsoline/villine	Der f 16	-	35	Kawamoto <i>et al.</i> , 2002
17	Protéines fixatrices de calcium	Der f 17	-	35	Tategaki <i>et al.</i> , 2000
18	Chitinase	Der p 18	-	54-63	O'Neil <i>et al.</i> , 2006
		Der f 18	-		Weber <i>et al.</i> , 2003
19	Peptide anti-microbien	Blot19	-	10	WHO/IUIS/GenBank;
20	Arginine kinase	Der p 20	-	?	WHO/IUIS/GenBank
21	Fonction inconnue	Der p 21	-	?	Weghofer <i>et al.</i> , 2008c
		Blo t 21	-		Gao <i>et al.</i> , 2007
22	Fonction inconnue	Der f 22	-	?	WHO/IUIS/GenBank
23	Fonction inconnue	Der p 23	-	?	WHO/IUIS/GenBank
24	Troponine C	Tyr p 24	-	?	Jeong <i>et al.</i> , 2010

WHO/IUIS/GenBank=Données non publiées mais séquences disponibles dans la liste des allergènes de WHO/IUIS ou dans GenBank

La majorité de ces allergènes d'acariens ont une activité enzymatique. Ce sont soit des cystéine-protéases (allergènes du groupe 1), des sérine-protéases ou trypsines (allergènes du groupe 3), des chymotrypsines (allergènes du groupe 6), des collagénases (allergènes du groupe 9), des amylases (allergènes du groupe 4), ou des chitinases (allergènes de groupes 15 et 18) qui correspondent aux enzymes digestives protéolytiques que l'acarien sécrète afin de dégrader les squames animales et les aliments protéiques nécessaires pour sa prolifération. Les allergènes du groupe 2, 13 et 14 sont des transporteurs de ligands. Ceux des groupes 10 et 11 sont des protéines musculaires (tropomyosine et paramyosine, respectivement), alors que d'autres allergènes ont une fonction qui est à l'heure actuelle inconnue.

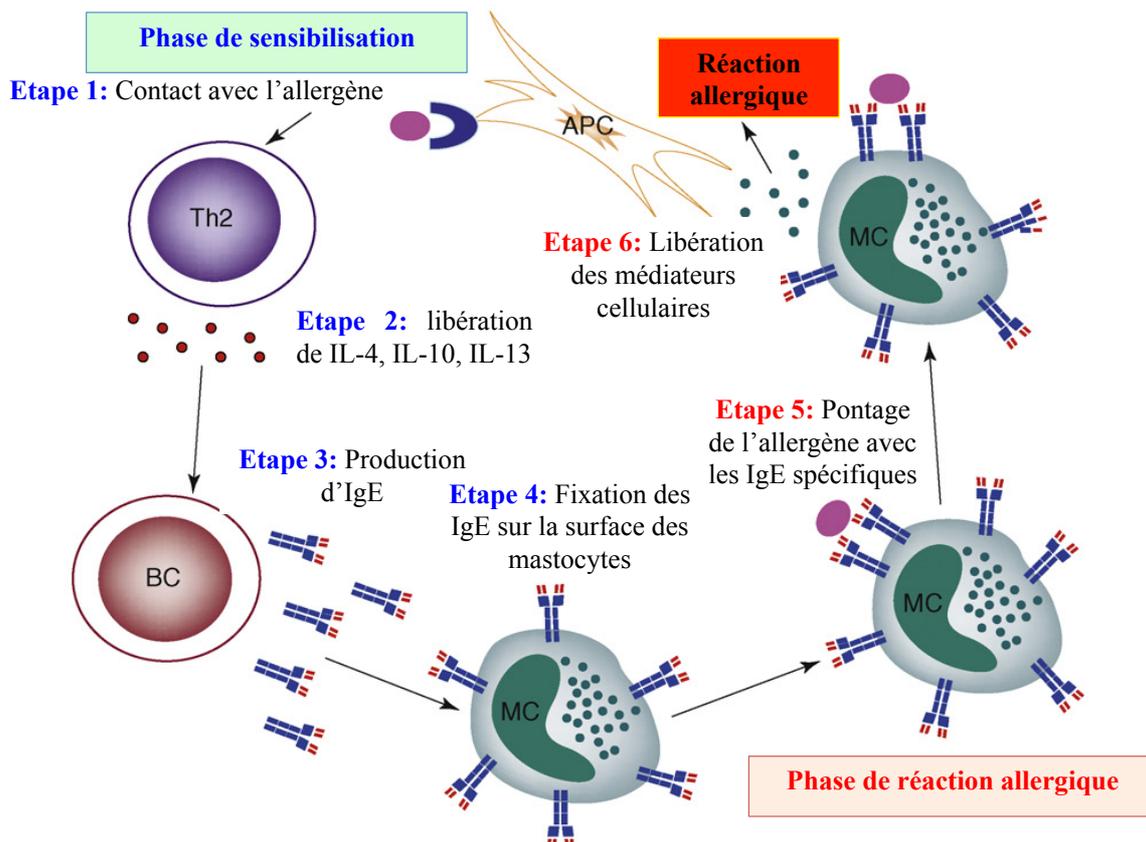
L'allergénicité de ces allergènes d'acariens est tout d'abord associée à leur propriété de liaison avec les anticorps "IgE". Dans ce contexte, les allergènes appelés "majeurs" sont ceux qui déclenchent une réaction allergique chez 50% ou plus des patients allergiques testés ou qui présentent des réponses immédiates à des tests cutanés chez au moins 90% des sujets ayant la maladie allergique en relation avec cet allergène (Chapman, 2008). Les allergènes majeurs d'acariens sont surtout ceux des groupes 1 et 2 (Der p 1 et Der p 2 isolés de *Dermatophagoides pteronyssinus* et Der f 1 et Der f 2 isolés de *Dermatophagoides farinae*), viennent par la suite les allergènes des groupes 4, 5 et 7 avec une allergénicité moyenne. Tandis que l'IgE-réactivité des autres allergènes (les symptômes rapportés et/ou observés lors du contact avec ces allergènes) est faible (tableau n°2) (Hales et al., 2006; Jacquet, 2011a). Toutefois, l'évaluation des taux de fixation des IgE avec les allergènes isolés de *Blomia tropicalis* (Blo t 2, Blo t 3, Blo t 5, Blo t 10, Blo t 11, Blo t 12 and Blo t 13) montre que l'allergène du groupe 5 "Blo t 5" est l'allergène majeur et le plus important de cette espèce d'acarien (Tsai et al., 2003; Chew et al., 2009).

### **I.2.3 Mécanisme d'activation du système immunitaire par les allergènes d'acariens**

Les allergènes induisent une réponse immunitaire adaptative humorale par la production des immunoglobulines (IgE) spécifiques. Le mécanisme par lequel l'allergène déclenche ces réactions passe par deux phases: la sensibilisation et la réaction allergique proprement dite (figure n°1).

Durant la première phase, l'allergène traverse la barrière épithéliale et entre en contact avec les cellules dendritiques (figure n°1, étape 1). Ces dernières présentent le peptide antigénique de l'allergène aux lymphocytes T naïfs (T<sub>H0</sub> CD4+) qui vont se différencier, chez les patients

allergiques, en cellules T helper 2 spécifiques ( $T_H2$ ) qui, à leur tour, sécrètent les cytokines (interleukine (IL)-4, IL-10 et IL-13) nécessaires à l'activation des lymphocytes B pour produire les anticorps "IgE-spécifiques" (figure n°1, étapes 2 et 3). Les IgE circulants se fixent alors à la surface des mastocytes et des basophiles (figure n°1, étape 4) (Shakib et *al.*, 2008; Galli et *al.*, 2008).



**Figure n°1:** Voie classique de la réponse allergique (adapté de Shakib et *al.*, 2008). APC, Cellules présentatrices d'antigènes (*Antigen Presenting Cells*); Th2, Lymphocytes T helper 2; BC, Cellules lymphocytaires B; MC, Cellules Mastocytaires.

Lors d'une exposition ultérieure au même allergène (phase de réaction allergique), ce dernier se lie aux molécules IgE spécifiques fixées à la surface des mastocytes et des basophiles grâce à ses déterminants antigéniques (épitopes) en créant ce qu'on appelle un pontage entre deux IgE adjacents et l'allergène (figure n°1, étape 5). Ce pontage stimule, par voie de signalisation, la dégranulation de ces mastocytes et la libération de médiateurs cellulaires tels que l'histamine et les leucotriènes (des lipides ayant un rôle chimiotactique) (figure n°1, étape 6). Ces médiateurs vont provoquer les symptômes de l'asthme, y compris la

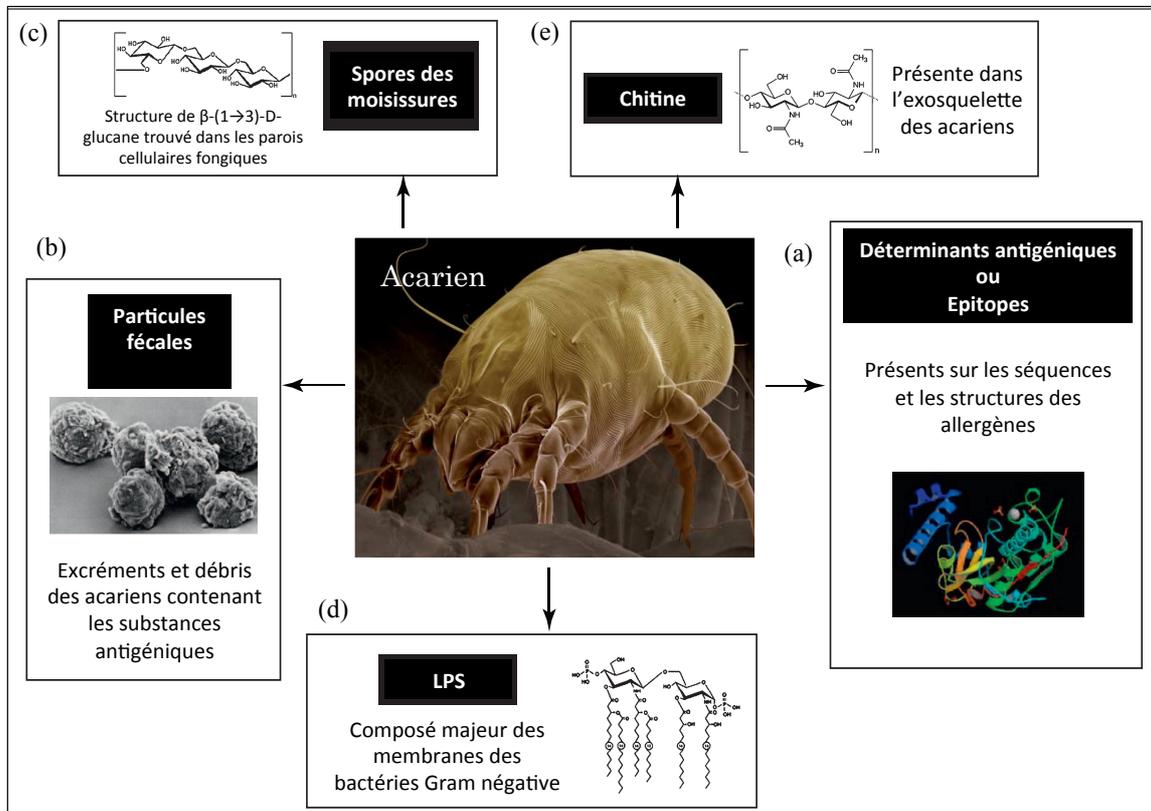
bronchoconstriction, la perméabilité vasculaire, la production de mucus et la vasodilatation (Galli et al., 2008).

Chez les sujets non allergiques, les lymphocytes T naïfs sont plutôt différenciés en cellules T<sub>H1</sub> qui produisent principalement de l'interféron- $\gamma$  (IFN- $\gamma$ ), de l'IL-2 et du TNF- $\beta$  (le facteur de nécrose tumorale- $\beta$ ). Ces cytokines inhibent la production des IgE en agissant directement sur les lymphocytes B.

En plus de l'activation des lymphocytes T<sub>H2</sub> et la production des IgE spécifiques, on assiste chez les sujets atopiques à une exacerbation des symptômes de l'asthme suite à l'activation des lymphocytes T<sub>H1</sub> et à d'autres lymphocytes T, notamment les T<sub>H17</sub> (caractérisés par la production des IL-17) et les lymphocytes T régulateurs (Treg) (Ouyang et al., 2008; Wang & Liu 2008; Zhao et al., 2012). Ces derniers ne sont pas des effecteurs de la réaction inflammatoire allergique (contrairement à T<sub>H1</sub>, T<sub>H2</sub> et T<sub>H17</sub>), mais des régulateurs faisant défaut chez les sujets allergiques (Lewkowich et al., 2005).

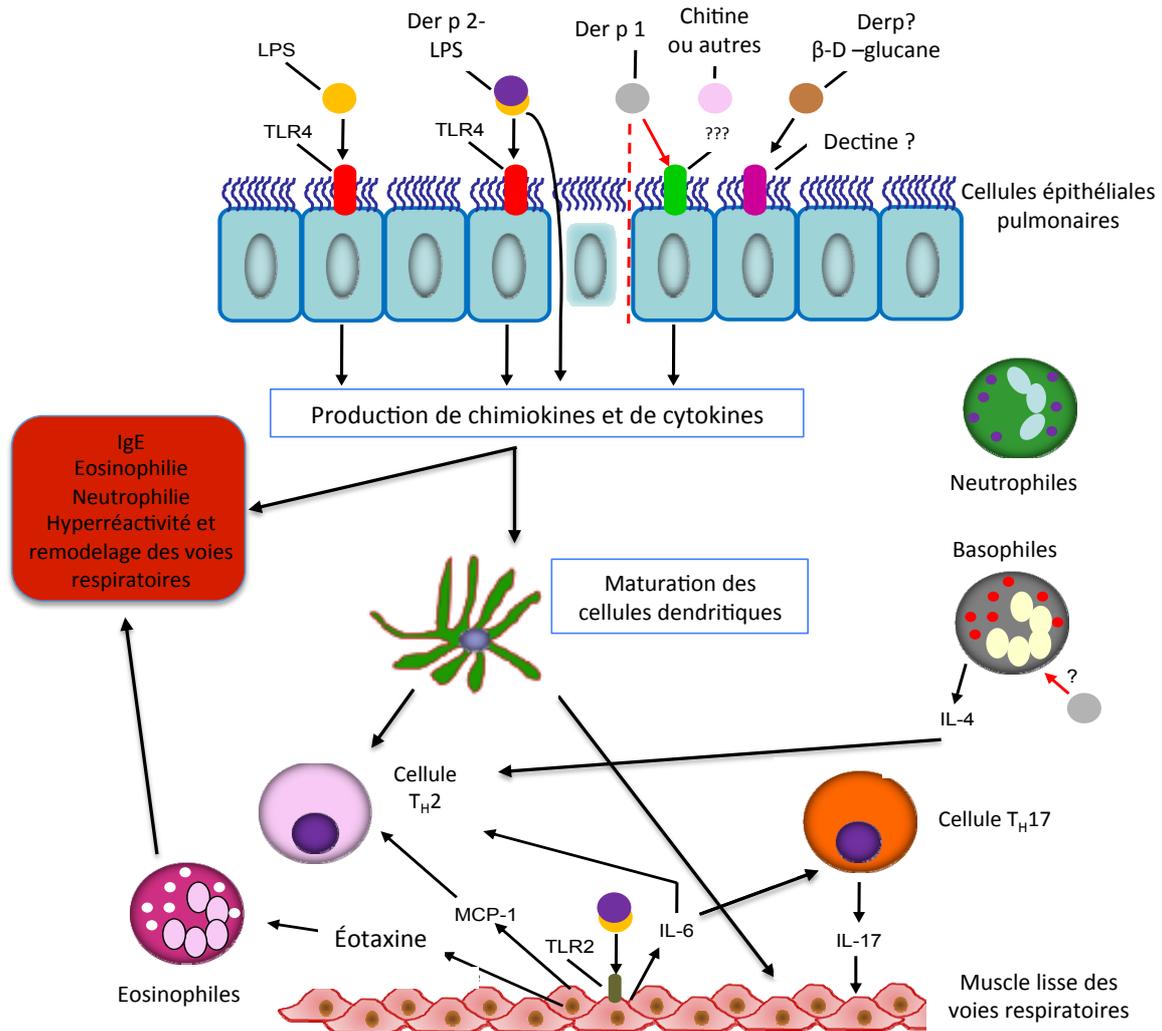
Cependant, en plus de ces réactions immunitaires, les allergènes d'acariens peuvent déclencher la réponse immunitaire innée et aggraver l'allergie grâce à leurs propriétés biochimiques et à leurs fonctions biologiques (Jacquet, 2009; Wills-Karp, 2010; Jacquet et al., 2011a et 2011b; Platts-Mills & Woodfolk, 2011). Par conséquent, dans le cas d'une réponse immunitaire de type allergique, l'activation du système immunitaire peut se faire par (1) l'allergène en lui-même *via* la fixation aux IgE spécifiques (suivant les étapes décrites ci-dessus) (figure n°2a, b) et (2) par des adjuvants ou immunomodulateurs présents dans l'environnement qui conjointement avec un antigène stimulent et renforcent le système immunitaire en activant la réponse immunitaire innée (figure n°2c, d, e).

Effectivement, des molécules microbiennes ayant le rôle d'adjuvant sont souvent présentes dans les poussières en association avec les allergènes d'acariens (figure n°2c, d, e). Plusieurs études ont mis en évidence la relation entre le  $\beta$ -D-glucane (constituants des parois des cellules fongiques), les endotoxines (constituants des membranes des bactéries Gram négatives) et la chitine (polymère à base de glucosamine qui forme l'exosquelette des acariens ainsi que des parois des cellules fongiques) et l'exacerbation des symptômes de l'allergie (Fahlbusch et al., 2003; Da silva et al., 2010; Jacquet, 2011a et 2011b).



**Figure n°2:** Molécules intervenant dans l'activation du système immunitaire en association avec les acariens (Issu de Gregory & Lloyd, 2011).

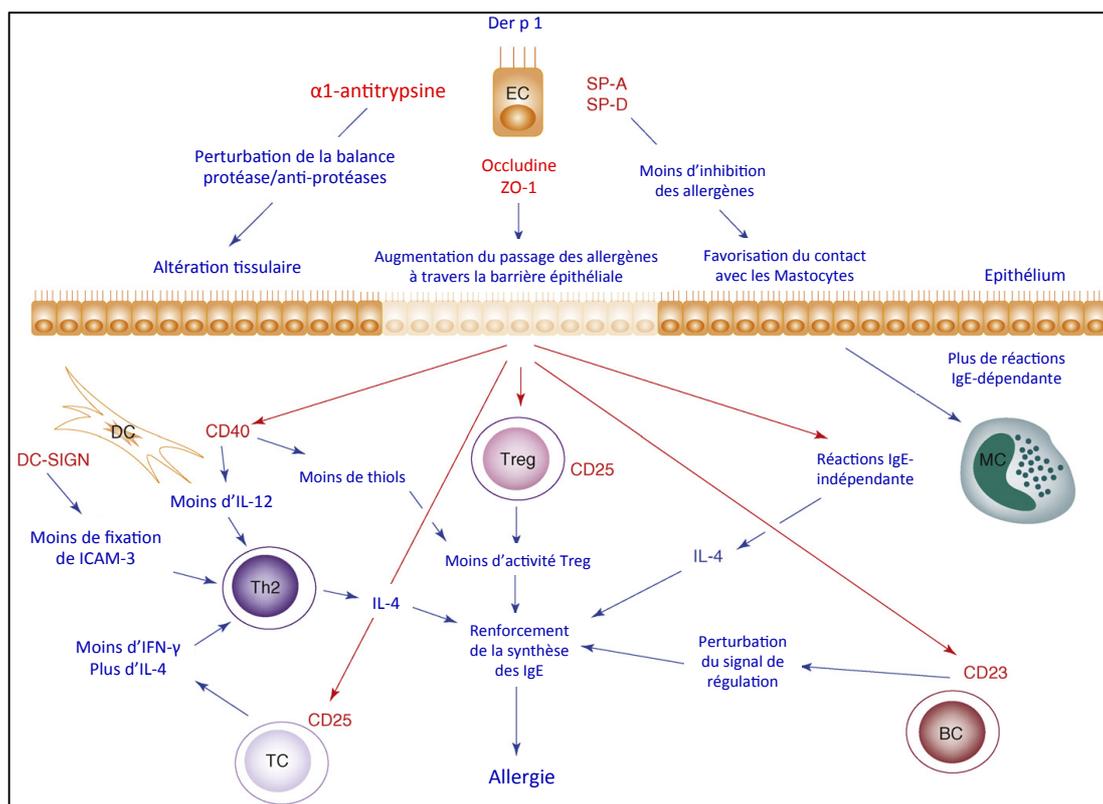
Parmi tous les allergènes d'acariens, seulement deux familles majeures de protéines ont été étudiées profondément et leur rôle est crucial dans la stimulation et l'activation de la réponse immunitaire innée: les protéases des groupes 1, 3, 6 et 9 par leur protéolyse cellulaire et les protéines fixatrices de ligands présentés par les groupes 2 et 7 et probablement 13, 14, 5 et 21, à travers leur transport facilité des adjuvants microbiens (Epton *et al.*, 2001; Mueller *et al.*, 2010a et 2010b; Chan *et al.*, 2006; Chapman *et al.*, 2007; Shakib *et al.*, 2008; Jacquet *et al.*, 2011c; Tsai *et al.*, 2011; Thomas, 2012). Les exemples les plus illustrés sont le rôle de la cystéine protéase de Der p 1 et la fixation des lipopolysaccharides par Der p 2 (figure n°3).



**Figure n°3:** Voies d'activation du système immunitaire par les allergènes d'acariens et leurs adjuvants (adapté de Jacquet, 2009). IL= Interleukine, TLR= *Toll-like receptor* (récepteurs de type Toll-like); MCP-1=*Monocyte chemoattractant protein 1*; LPS= Lipopolysaccharides.

Le contact des adjuvants microbiens (les lipopolysaccharides, le glucane ou la chitine) avec les allergènes d'acariens peut activer les récepteurs sensibles comme les PRR (*Pattern Recognition Receptors*) trouvés à la surface des cellules épithéliales pulmonaires et induire la réponse immunitaire innée. Cette stimulation conduit à la production des chimiokines et des cytokines qui favorisent le recrutement des cellules leucocytaires du système immunitaire et provoquer, ainsi, une éosinophilie (augmentation permanente des éosinophiles), une neutrophilie (augmentation permanente des neutrophiles) et l'hyperréactivité et le remodelage des voies respiratoires.

La voie d'activation de cette réponse immunitaire par Der p 1 a été largement étudiée, et a fait l'objet de plusieurs travaux qui ont montré la complexité du mécanisme d'action de cet allergène avec son activité enzymatique liée à sa fonction protéasique. Pour simplifier le rôle de cette activité, nous avons repris la figure n°4 adaptée de Shakib *et al.* (2008) où les cellules cibles effectrices et leurs conséquences immunologiques sont pointées par des flèches rouges et bleues, respectivement.



**Figure n°4:** Mécanisme d'action de l'allergène majeur Der p 1 (adapté de Shakib *et al.*, 2008).

EC= Cellules épithéliales; BC= Cellules lymphocytaires B; TC= Cellules Lymphocytaires T; Treg= Lymphocytes T régulateurs; MC= Cellules Mastocytaires; SP-A, SP-D= Surfactant pulmonaire A et D; ZO-1= *Zonula Occludens-1*; DC-SIGN= Molécules d'adhésion intracellulaire non-intégrine 3 spécifiques aux cellules dendritiques; ICAM-3= Molécule d'adhésion intercellulaire 3.

Der p 1 est capable de modifier la perméabilité de l'épithélium bronchique et les jonctions serrées en clivant le domaine extracellulaire de l'occludine et des claudines associées aux protéines intracellulaires *Zona Occludens* (ZO) (Wan *et al.*, 2000), ce qui facilite le passage de cet allergène (ainsi que d'autres) *via* les cellules épithéliales et facilite leur accessibilité aux cellules dendritiques. Cette activité est également accomplie par les allergènes des groupes 3, 6 et 9 (Der p 3, Der p 6 et Der p 9) (Tomee *et al.*, 1998; King *et al.*, 1996).

Adam *et al.* (2006) ont montré que Der p 1 est capable de stimuler la production des IL-6, des IL-8, des éotaxines (des chémokines impliquées dans le recrutement des éosinophiles) et des GM-CSF (facteurs destinés à stimuler la multiplication des granulocytes et des macrophages) mais par une voie d'activation PAR-2 (récepteurs activés par des tryptase) indépendante, contrairement à Der p 3 et à Der p 9. Par conséquent, la présence d'IL-6 induit la polarisation des cellules T<sub>H0</sub> en cellules T<sub>H17</sub> qui accentuent la pathologie des voies respiratoires en agissant sur les muscles lisses des voies aériennes (Wang & Liu, 2008). En plus, Der p 1 inactive catalytiquement la protéine  $\alpha$ 1-antitrypsine (protéine anti-protéase) qui protège la muqueuse respiratoire contre les sérine-protéases (Kalsheker *et al.*, 1996). Der p 1 et Der f 1 peuvent, également, inhiber les protéines A et D (SP-A et SP-D) du surfactant pulmonaire ayant un rôle inhibiteur sur la fixation de ces allergènes sur les IgE liés à la surface des mastocytes afin d'empêcher la dégranulation de ces derniers (Deb *et al.*, 2007).

Der p 1 peut, également, avoir une action sur les différentes cellules leucocytaires. En effet, il peut induire la sécrétion de l'IL-4 par les mastocytes et les basophiles indépendamment d'une sensibilisation IgE-dépendante (Machado *et al.*, 1996). De plus, il est susceptible d'agir sur la surface des lymphocytes B en clivant spécifiquement les récepteurs de faible affinité CD23 des IgE et par conséquent provoquer la suppression du mécanisme de la régulation de production des ces anticorps (Schulz *et al.*, 1997). Cet allergène clive également les récepteurs CD25 des lymphocytes T provoquant ainsi une diminution de la sécrétion de l'INT- $\gamma$  et donc l'orientation de la prolifération des lymphocytes T vers une réponse T<sub>H2</sub>. Der p 1 a également cette même action sur la surface des cellules dendritiques en éliminant leurs récepteurs CD40 et réduisant ainsi la production des IL-12 qui jouent un rôle dans la prolifération des cellules vers un profil T<sub>H1</sub> (Ghaemmaghami *et al.*, 2002). Le clivage des molécules DC-SIGN par Der p 1 peut de même favoriser la voie T<sub>H2</sub> (Furmonaviciene *et al.*, 2007). A côté de toutes ces actions, Fuiano & Incorvaia (2012) ont aussi signalé l'importance du rôle de Der p 1 dans le développement de la dermatite atopique chez les enfants.

D'autre part, la fonction des allergènes du groupe 2 "Der p 2" et "Der f 2" était inconnue jusqu'à ce que l'on découvre leur similarité structurale avec la molécule MD-2 (*Myeloid Differentiation 2*) qui lie les lipopolysaccharides des bactéries gram négatives et déclenche la voie de signalisation de type TLR-4 (*Toll-like receptor 4*) dépendante en se fixant sur ces récepteurs trouvés à la surface des cellules épithéliales pulmonaires, des macrophages et des cellules dendritiques (Ichikawa *et al.*, 1998; Derewenda *et al.*, 2002; Suzuki *et al.*, 2006; Kim

et *al.*, 2007). De nombreux allergènes sont membres de la famille de protéines fixatrices de lipides, ce qui suggère que leur activité liée aux lipides associés pourraient également contribuer à l'allergénicité de ces allergènes (Trompette et *al.*, 2009). Der p 2 peut, également, affecter les cellules musculaires lisses en stimulant une voie de signalisation MyD88 dépendante du récepteur TLR-2 (Chiou & Lin, 2009). Récemment, Tsai et *al.* (2011) ont pu démontrer qu'en plus de l'induction de la réponse immunitaire TLR-4 dépendante, Der p 2 peut aussi activer les cellules B et induire la production des cytokines IL-1b, CXCL10, IL-8 et TNF- $\alpha$ .

Peu d'informations sont trouvées sur les allergènes de groupe 7. Mueller et *al.* (2010b) ont classé cet allergène Der p 7 parmi les protéines fixatrices de petits ligands, en démontrant qu'il fixe la polymyxine B, peptide isolé de *Bacillus polymyxa*, mais pas les lipopolysaccharides. Cet allergène pourrait donc déclencher une réponse immunitaire en activant le récepteur TLR-2 (Jaquet et *al.*, 2011a).

Le mécanisme des allergènes du groupe 5 est inconnu, mais des travaux ont montré que l'allergène recombinant de Der p 5 stimule la production des IL-6 et IL-8 avec un mécanisme protéase-indépendant (Kauffman et *al.*, 2006). Par ailleurs, la structure tridimensionnelle de Der p 5 présente une cavité hydrophobe qui pourrait accueillir et transporter des ligands hydrophobes et ainsi déclencher une réponse immunitaire comme celle des allergènes des groupes 2 ou 7 (Mueller et *al.*, 2010a; Thomas, 2012).

Nous voyons clairement que les propriétés biochimiques de ces allergènes contribuent directement à l'aggravation de l'allergie et que la fonction biologique de ces allergènes joue donc un rôle primordial dans la compréhension des mécanismes d'activation du système immunitaire.

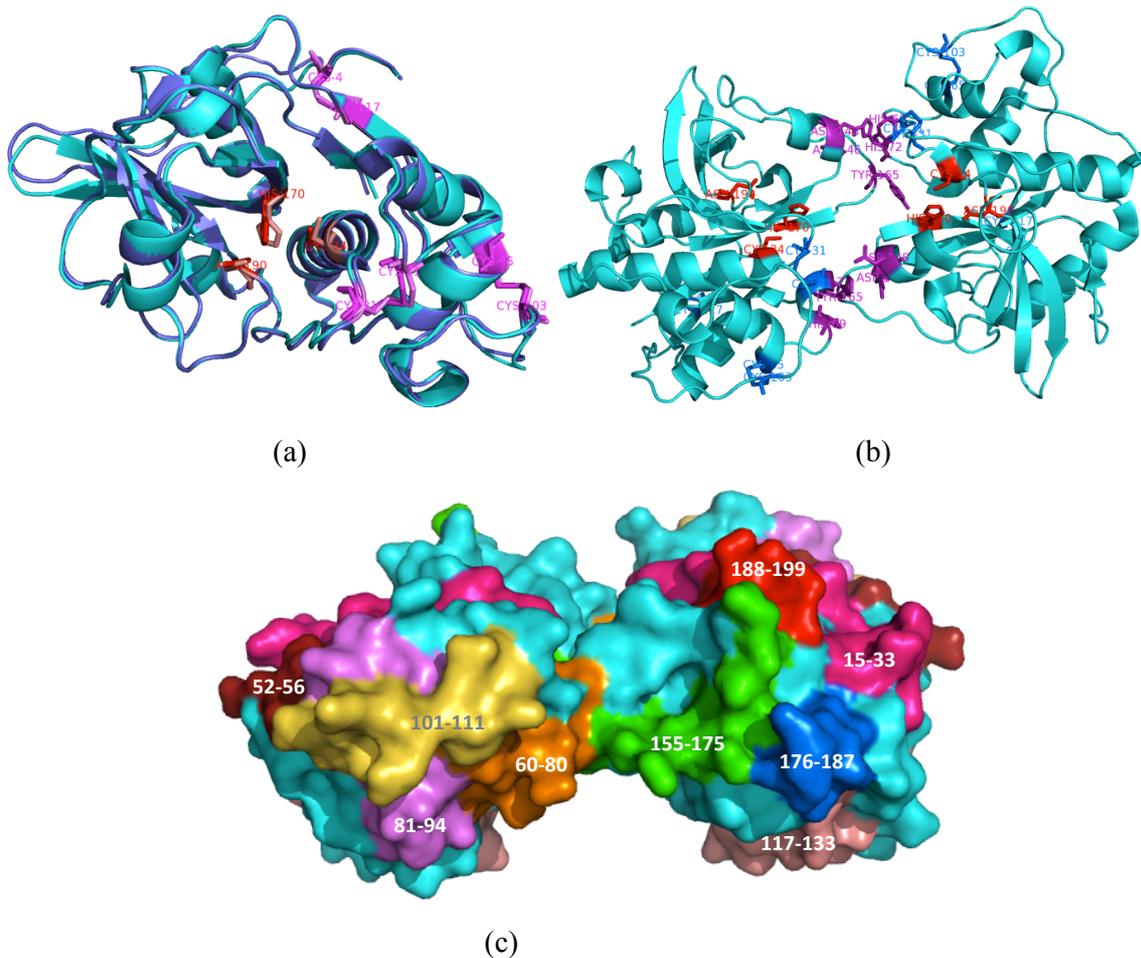
#### **I.2.4 Rôle de la structure des allergènes dans leur allergénicité**

La structure 3D des allergènes définit en partie leur fonction biologique (Pomés, 2002). Ainsi, la forme allongée de la tropomyosine de *Dermatophagoides pteronyssinus* (Der p 10) est appropriée pour la contraction musculaire, alors que la forme globulaire de Der p 2 avec sa cavité hydrophobe lui permet de transporter des ligands hydrophobes tels que les LPS. Cette structure 3D pourrait nous fournir des informations quant à la fonction des allergènes, nous donner une idée sur la position des épitopes et, ainsi, contribuer à l'étude des mécanismes de

leur allergénicité et permettre la conception de variants hypoallergéniques pour l'immunothérapie spécifique.

Les allergènes du groupe 1 sont des cystéines protéases et leurs structures 3D ressemblent à celles des papaines (papain-like). Les structures de Der p 1 et de Der f 1 ont été résolues expérimentalement par cristallographie aux rayons-X (code PDB 2AS8 et 3DS6, respectivement) (de Halleux et *al.*, 2006; Chruszcz et *al.*, 2009). Ces structures ont une forme globulaire avec deux domaines appelés "domaine *Left*" (constitué majoritairement d'hélices  $\alpha$  formées par les résidus 101 à 196, selon la numérotation du fichier PDB 2AS8 de Der p 1) et le "domaine *Right*" (constitué d'hélices  $\alpha$  et de feuillets  $\beta$  formés par les résidus 197 à 299, selon la numérotation du fichier PDB 2AS8 de Der p 1). Ces deux domaines délimitent un sillon comportant la triade catalytique C34-H170-N190 (selon la numérotation du fichier PDB 2AS8 de Der p 1) (figure n°5). Trois ponts disulfures sont présents dans Der p 1 et conservés dans Der f 1 (C4-C117, C31-C71, et C65-C103, selon la numérotation du fichier PDB 2AS8). Toutefois, la structure de Der p 1 est susceptible de dimériser contrairement à ce qui a été trouvé pour Der f 1 (De Halleux et *al.*, 2006). L'interface du dimère de Der p 1 est entre autres formée de trois paires de résidus polaires et aromatiques (H69-H72, D146-D148 et Q166-N168, selon la numérotation du fichier PDB 2AS8) qui jouent un rôle crucial dans la stabilité de ce dimère. Les structures de Der p 1 et de Der f 1 montrent des différences au niveau de quatre régions de surface et dans le site de fixation de l'ion manganèse mais une conservation de certaines zones qui sont probablement responsables de la réactivité croisée entre ces deux allergènes (Chruszcz et *al.*, 2009).

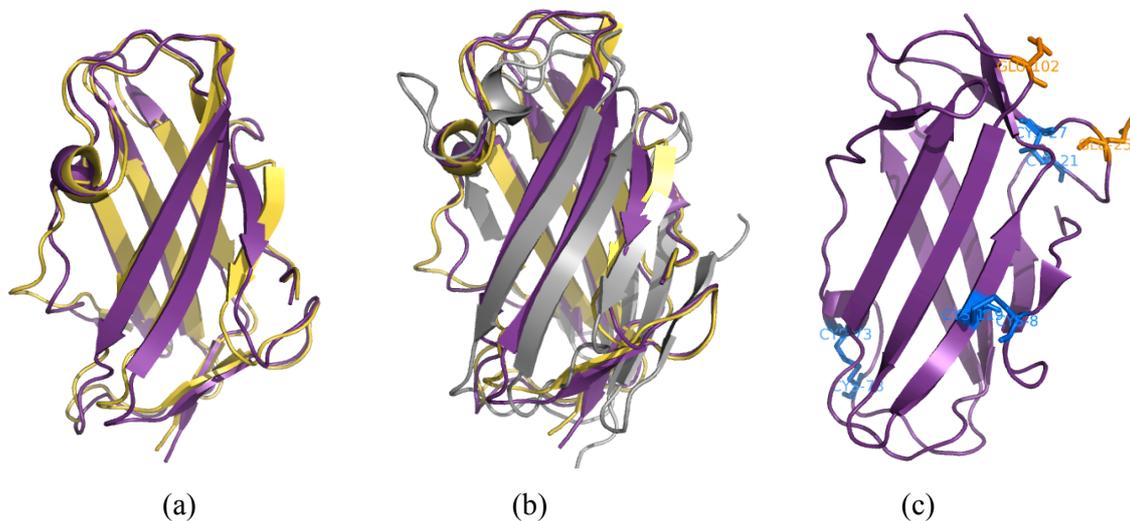
Par ailleurs, plusieurs études ont fait l'objet de la caractérisation des épitopes de ces allergènes du groupe 1. Greene & Thomas (1992) et Jeannin et *al.* (1992) ont identifié les segments 15-33, 60-80, 81-94, 101-111, 117-133, 155-175, 176-187 et 188-199, comme épitopes dans Der p 1 (figure n°5c). Parmi ces épitopes, seulement celui en position 60-80 est conservé dans Der f 1, expliquant ainsi la réactivité croisée entre ces deux allergènes (Collins et *al.*, 1996). Récemment, plus de deux structures d'anticorps Fab-Der p 1 et Fab-Der f 1 ont été résolues (Chruszcz et *al.*, 2009; Zhong et *al.*, 2011; Chruszcz et *al.*, 2012). Ces structures fournissent plus d'informations sur la réactivité croisée entre Der p 1 et Der f 1.



**Figure n°5:** Structures tridimensionnelles des allergènes d'acariens du groupe 1. (a) Superposition de structure entre Der p 1 (en cyan, code PDB 2AS8) et Der f 1 (en bleu, code PDB 3D6S). (b) Structure dimérique de Der p 1 (code PDB 2AS8), les résidus qui stabilisent l'homodimère sont présentés en bleu. Les résidus catalytiques et les ponts disulfures sont présentés en sticks rouges et en magenta, respectivement. (c) Présentation en surface de l'homodimère de Der p 1 en cyan et la position des différents épitopes déterminés expérimentalement.

Les structures des deux allergènes majeurs Der p 2 et Der f 2 ont été également résolues expérimentalement par cristallographie aux rayons X (code PDB 1KTJ et 2F08, respectivement) (Derewenda *et al.*, 2002; Suzuki *et al.*, 2006). Ces deux allergènes montrent une grande similarité structurale avec la molécule MD-2 qui joue un rôle dans la reconnaissance des LPS et dans leur présentation aux récepteurs TLR-4 (figure n°6a et b). Leurs structures sont constituées principalement par des feuillettes  $\beta$  antiparallèles formant ce qu'on appelle des " $\beta$ -sandwich". Trois ponts disulfures sont trouvés dans Der p 2 (C21-C27,

C73-C78 et C8-C119, selon la numérotation dans le fichier PDB 1KTJ) (Derewenda et *al.*, 2002) (figure n°6c). Ces ponts disulfures qui stabilisent la structure sont également conservés dans Der f 2 (Ichikawa et *al.*, 1998). Le repliement de ces structures forme une grande cavité hydrophobe qui fixe les LPS (Ichikawa et *al.*, 2009). La cavité hydrophobe de Der f 2 a une surface moléculaire de 413 Å<sup>2</sup> et un volume de 348 Å<sup>3</sup> (Suzuki et *al.*, 2006). Cette cavité est nettement inférieure à celle de Der p 2 (avec une surface de 1150,6 Å<sup>2</sup> et un volume de 1509 Å<sup>3</sup>) et à celle de MD-2 (avec une surface de 2374,1 Å<sup>2</sup> et un volume 1099,4 Å<sup>3</sup>). Cependant, Ichikawa et *al.* (2009) ont montré que Der f 2 lie les LPS entre ses deux grands feuillets β, d'une manière similaire à MD-2, avec une affinité de l'ordre du nanomolaire. Cette conservation de structure pourrait donc être l'un des facteurs les plus cruciaux qui déterminerait l'allergénicité de ces allergènes du groupe 2.

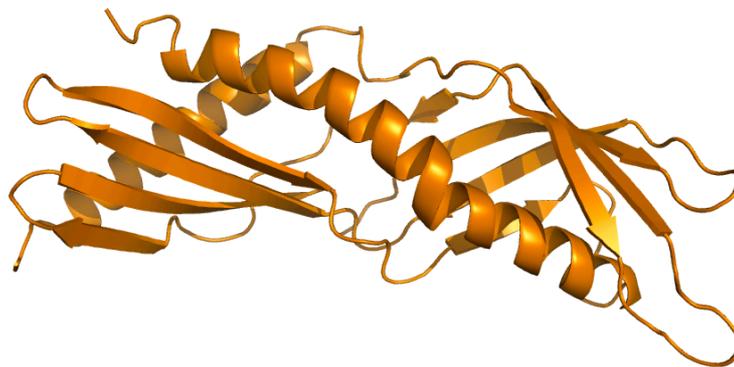


**Figure n°6:** Structures tridimensionnelles des allergènes d'acariens du groupe 2. (a) Superposition de structure entre Der p 2 (en violet, code PDB 1KTJ) et Der f 2 (en jaune, code PDB 2F08). (b) superposition de structure de la molécule MD-2 (en gris, code PDB 3FXI) avec Der p 2 et Der f 2. (c) Position des ponts disulfures (en sticks bleus) et les résidus épitopes majeurs E25 et E102 (en sticks oranges) dans Der p 2.

De plus, plusieurs épitopes ont été identifiés pour Der p 2 (Smith & Chapman, 1996; Hakkaart et *al.*, 1998; Mueller et *al.*, 2001; Derewenda et *al.*, 2002; Ipsen et *al.*, 2004.). Chaque étude a identifié différents résidus jouant un rôle important dans la fixation avec les IgE mais les mutants générés n'ont pas été efficaces en tant que candidats pour des vaccins pour l'immunothérapie spécifique à Der p 2 sauf pour les deux résidus E25 et E102 (figure n°6c) (Reginald, 2006).

Il faut, par ailleurs, souligner que les deux allergènes majeurs Der p 1 et Der p 2 ont, respectivement, 23 et 13 isoformes. L'analyse de la substitution d'acides aminés dans ces derniers peut avoir un rôle dans la réactivité allergique et dans l'allergénicité de ces allergènes si ces substitutions se trouvent dans leurs épitopes B ou T (Chapman et *al.*, 2009).

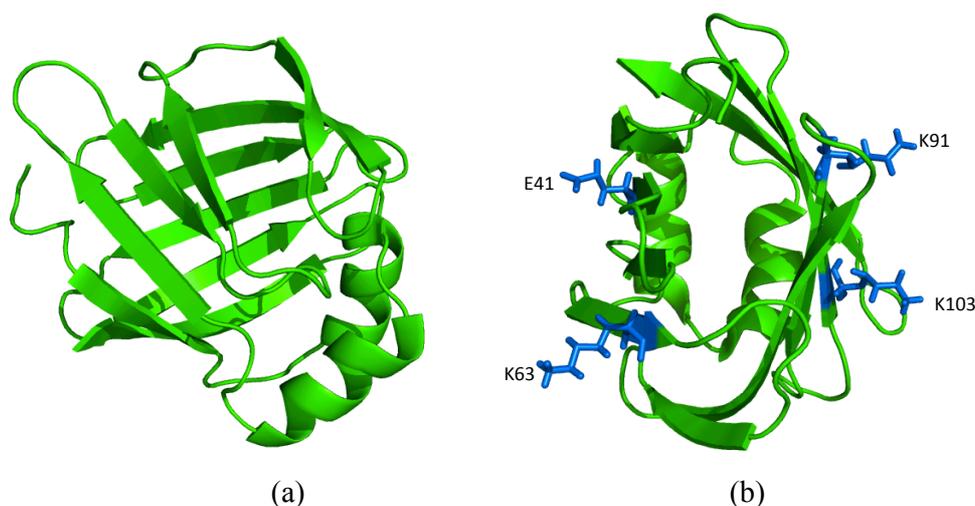
Les allergènes des groupes 7 et 13 sont également considérés comme des protéines fixatrices de ligands hydrophobes. Parmi les allergènes de cette famille 7, seule la structure de Der p 7 est résolue expérimentalement par RMN mais ses épitopes n'ont pas été identifiés (Mueller et *al.*, 2010b). Cet allergène présente une structure allongée formée d'hélices  $\alpha$  et de feuillets  $\beta$ , avec de petites cavités hydrophobes au centre de la molécule (figure n°7). La forme et la taille de ces cavités peuvent expliquer le fait que cet allergène ne puisse pas lier les LPS mais lie des petites molécules comme la polymyxine B (Mueller et *al.*, 2010b).



**Figure n°7:** Structure tridimensionnelle cristallographique de Der p 7 (code PDB 3H4Z).

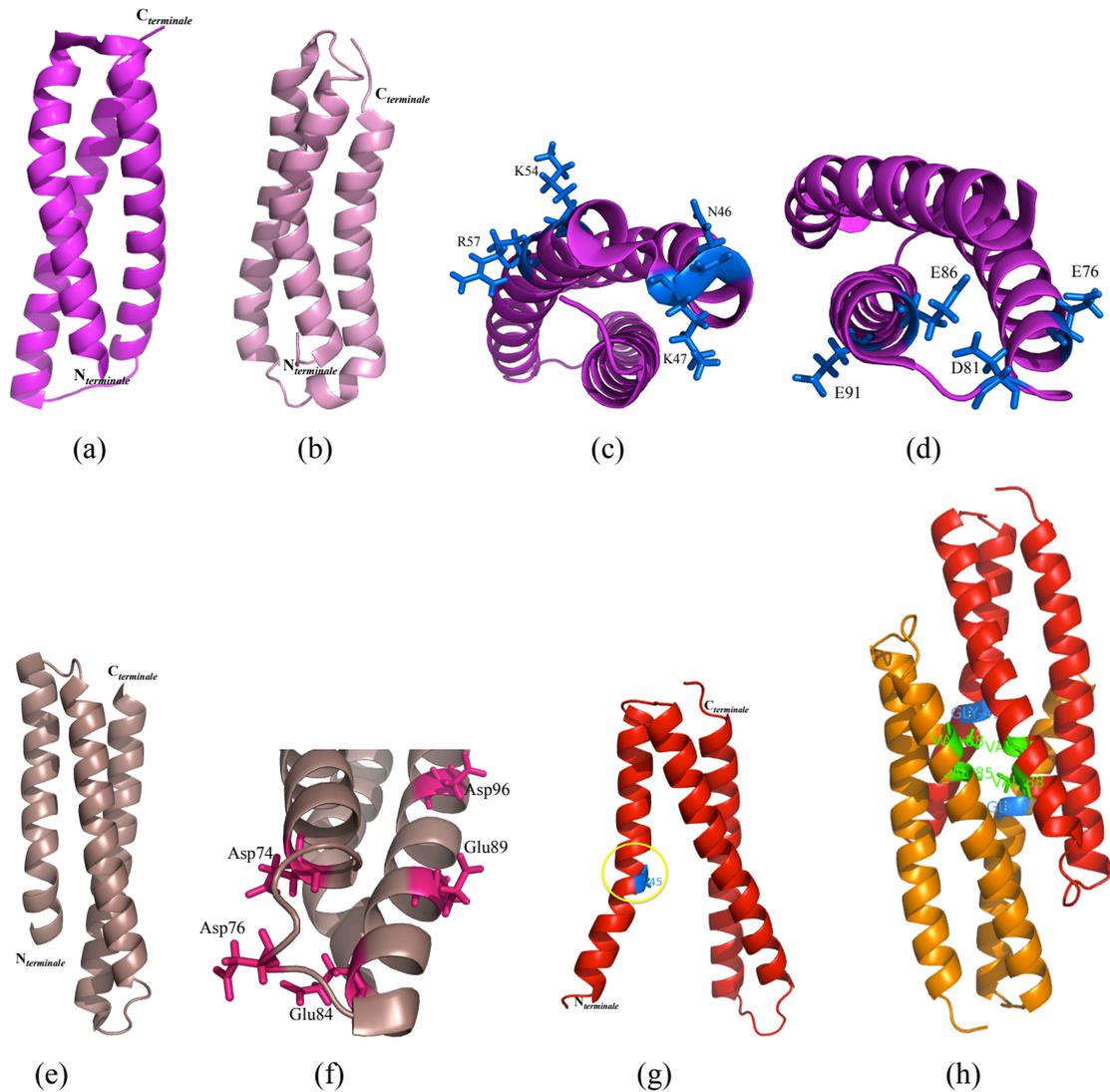
Par ailleurs, Shen et *al.* (2011) et Chou et *al.* (2011) ont créé, *in silico*, un modèle de structure pour Der f 7 et ont pu définir, par la méthode des peptides contigus se chevauchant (*Contiguous Overlapping Peptides-COP*), deux régions épitopes (S56-I57-L58-D59-P60 et G46-I47-L48-D49-F50, selon la numérotation dans le fichier PDB 3H4Z de Der p 7) responsables de la fixation des IgE et jouant un rôle sur la réactivité croisée entre Der f 7 et Der p 7. Bien que les premiers essais de cristallisation de Der f 7 aient donné de bons résultats, la structure de cet allergène n'est pas encore complètement déterminée (Tan et *al.*, 2011).

La structure RMN de l'allergène Der f 13 a été déterminée par Chan et *al.* (2006) (code PDB: 2A0A). Cette structure ressemble beaucoup aux protéines fixatrices d'acides gras (*fatty-acid-binding proteins -FABPs*). Le feuillet  $\beta$  de la partie *N\_terminale* est suivi par un motif hélice-coude-hélice et puis neuf autres brins  $\beta$  antiparallèles (figure n°8a). Un seul épitope B composé des quatre acides aminés chargés E41, K63, K91 et K103 (selon la numérotation dans le fichier PDB 2A0A) a été caractérisé pour cet allergène (figure n°8b) (Chan et *al.*, 2006).



**Figure n°8:** Structure RMN de Der f 13 (code PDB 2A02) (a) et position de ses épitopes en sticks bleus (b).

Les allergènes du groupe 5 sont également importants. Seules les structures des allergènes majeurs Blo t 5 et Der p 5 ont été déterminées expérimentalement par RMN et par RX, respectivement (Naik et *al.*, 2008; Chan et *al.*, 2008; Mueller et *al.*, 2010a). Leurs structures sont composées de trois hélices alpha formant une organisation en *coiled-coil*. Deux structures ont été résolues pour Blo t 5: code PDB 2JMH (Naik et *al.*, 2008) et 2JRK (Chan et *al.*, 2008). Le repliement de cette protéine est différent entre ces deux structures (figure n°9a et b). Cependant, les interactions entre les hélices de la structure de Der p 5 ressemblent plus à la structure 2JMH de Blo t 5 (figure n°9a, b et g), ce qui plaide en faveur du PDB 2JMH en tant que structure de Blo t 5 (Mueller et *al.*, 2010a).



**Figure n°9:** Structures 3D des allergènes d'acariens des groupes 5 et 21. (a) Structure RMN de Blo t 5, code PDB 2JMH. (b) Structure RMN de Blo t 5, code PDB 2JRK. (c) Position des résidus majeurs de l'épitope conformationnel de Blo t 5 en sticks bleus. (d) Position des résidus majeurs de l'épitope linéaire de Blo t 5 en sticks bleus. (e) Structure RMN de Blo t 21, code PDB 2LM9. (f) Position des résidus majeurs de l'épitope conformationnel de Blo t 21 en sticks roses. (g) Chaîne A de la structure cristallographique de Der p 5 (code PDB 3MQ1; 3MQ1\_A), la G45 est représenté en stick bleu. (h) Structure de l'homodimère de Der p 5 (chaîne A en rouge et chaîne B en orange; 3MQ1\_AB): position des glycines 45 (en sticks bleus) et des valines 85 et 88 (en sticks verts).

Par ailleurs, la structure de Blo t 5 est rapportée sous forme de monomère (Chan et *al.*, 2008; Naik et *al.*, 2008) tandis que celle de Der p 5 (code PDB 3MQ1) a été trouvée comme étant un homodimère formant une grande cavité hydrophobe au centre de la molécule (de 1777,8 Å<sup>2</sup> de surface et 3051,2 Å<sup>3</sup> de volume) (figure n°9h) (Mueller et *al.*, 2010a). D'après ces auteurs, la glycine en position 45 (G45) de l'hélice N<sub>terminale</sub> pourrait jouer un rôle dans la dimérisation de Der p 5 (figure n°9g et h). Cette glycine pourrait être importante pour la flexibilité de cette hélice et conduire à la formation du dimère où les deux chaînes de Der p 5 sont enchevêtrées (figure n°9g).

La séquence de Der p 5 ne contient pas de résidus cystéines, mais sa structure dimérique semble être stabilisée par une liaison de type "valine zipper" entre la valine 85 de la première chaîne et la valine 88 de la deuxième chaîne (figure n°9h). Ainsi, Mueller et *al.* (2010a) ont suggéré que la création d'une cavité hydrophobe appuie l'hypothèse que cet allergène pourrait être considéré comme une protéine fixatrice de ligands, comme les allergènes du groupe 2, et contribuerait ainsi à l'aggravation des symptômes de l'allergie.

Deux épitopes B ont été identifiés dans Blo t 5, un conformationnel (Naik et *al.*, 2008) et l'autre linéaire (Chan et *al.*, 2008) (figure n°9c et d), mais aucun épitope n'est encore connu pour Der p 5 jusqu'à présent. L'épitope conformationnel de Blo t 5 est composé de deux résidus (N46 et K47) situés à l'extrémité de la deuxième hélice et de deux autres résidus (K54 et R57) dans le début de la troisième hélice (figure n°9c). L'épitope linéaire, quant à lui, est composé de 15 acides aminés contigus (E76-E91). Dans cet épitope, quatre résidus sont définis comme majeurs jouant un rôle crucial dans la fixation des IgE: E76, D81, E86 et E91 (figure n°9d).

Les allergènes de la famille 21 présentent des similarités de séquences avec leurs homologues de la famille 5. La structure 3D de l'allergène Der p 21 n'a pas été résolue expérimentalement, mais a été étudiée par diffusion des rayons-X aux petits angles (*Small Angle X-ray Scattering-SAXS*) et par dichroïsme circulaire. Les résultats de ces expériences suggèrent que cet allergène soit également composé de trois hélices alpha et pourrait présenter une structure sous forme de dimère avec une forme allongée. Les simulations *ab initio* réalisées proposent un mode de dimérisation différent que celui de Der p 5 (Weghofer et *al.*, 2008c). Cependant, la structure de Blo t 21 a été récemment déterminée expérimentalement par RMN (code PDB 2LM9, figure n°9e) (Tan et *al.*, 2012). Cette structure est sous forme de monomère avec une conformation très semblable à celle de Blo t 5. Ces auteurs (Tan et *al.*, 2012) ont également

identifié la position d'un épitope conformationnel de Blo t 21 composé des résidus E74, D79 et le E84 d'un côté, et de E89 et D96 dans l'autre côté (figure n°9f).

Les résultats obtenus dans toutes ces études fournissent plus d'informations sur les caractéristiques moléculaires et structurales relatives à l'allergénicité et à la réactivité croisée entre les allergènes. Ces informations pourront servir à la conception des hypoallergènes pour l'immunothérapie spécifique.

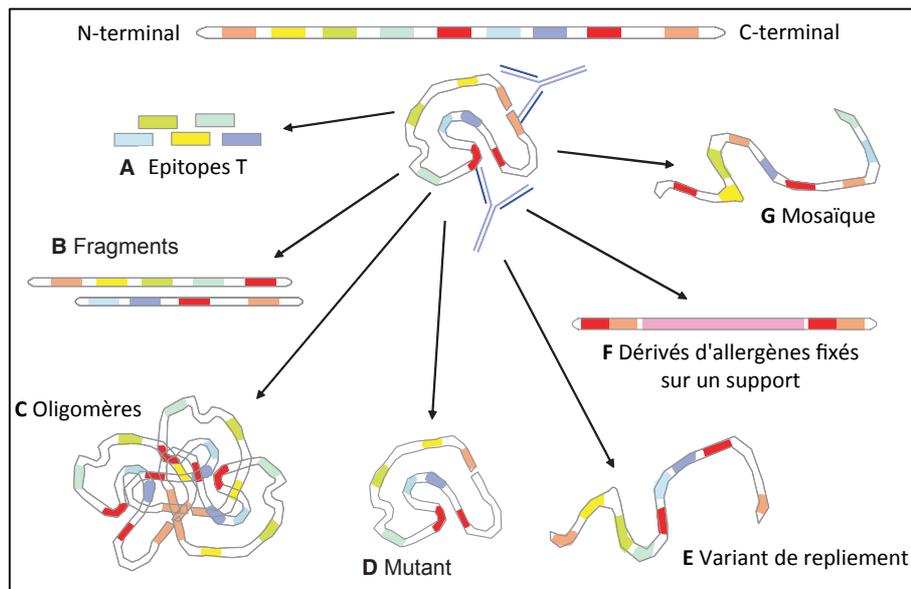
### I. 3 Traitement des allergies respiratoires par l'immunothérapie spécifique

L'immunothérapie spécifique ou la désensibilisation implique l'administration de concentrations croissantes d'extraits spécifiques de l'antigène aux patients allergiques dans le but d'induire un état de tolérance immunologique traduit par une diminution de l'hypersensibilité et une réduction des symptômes lors d'une exposition ultérieure à l'antigène (Durham, 2006). Elle rétablit l'équilibre des réponses  $T_H1/T_H2$ . En effet, l'immunothérapie avec des extraits de *Dermatophagoides pteronyssinus* améliore la fonction pulmonaire des enfants atteints d'asthme allergique. Cette amélioration est assistée par la diminution des taux sériques d'IL-4 et d'IL-13 (qui interviennent dans la réponse immunitaire  $T_H2$ ) et à l'augmentation des taux sériques d'IFN- $\gamma$  favorisant ainsi une réponse  $T_H1$  (Chen et al., 2009). Par inhibition de la fonction des lymphocytes  $T_H2$ , les lymphocytes  $T_H1$  peuvent avoir un effet protecteur et bénéfique dans l'asthme. Les cellules Treg sont également une cible prometteuse pour l'immunothérapie spécifique. Huang et al. (2011) ont montré que ces lymphocytes T régulateurs contribuent à l'atténuation de l'allergénicité de Blo t 5 par la production de l'IL-35. La caractérisation de l'allergénicité des allergènes d'acariens pourrait donc ouvrir la porte à de nouvelles stratégies thérapeutiques pour bloquer efficacement cette cascade inflammatoire exacerbée (Jacquet, 2011b).

La biologie moléculaire des allergènes a fait de grands progrès au cours de ces 20 dernières années au travers du clonage et du séquençage des allergènes, de l'expression des allergènes recombinants, de la détermination de leurs structures tridimensionnelles, et des essais cliniques de vaccins d'allergènes recombinants (Chapman et al., 2009; Pomés, 2010). La disponibilité des structures des allergènes majeurs les plus courants a contribué fortement à la compréhension de la fonction biologique et à la détermination des épitopes des allergènes: ce qui a ainsi permis de produire des vaccins antiallergiques qui permettent un ciblage spécifique des mécanismes des maladies allergiques. De plus, le développement des systèmes

d'expression des allergènes recombinants ont permis la conception des formes modifiées des allergènes appelées «hypoallergènes» dans lesquelles les épitopes B reconnus par les IgE sont éliminés tandis que les épitopes T et ceux reconnus par les IgG sont conservés, permettant ainsi l'induction des anticorps de type IgG neutralisants (Valenta, 2002; Linhart & Valenta, 2012).

Pour les allergènes inhalés, les épitopes conformationnels semblent être la principale cible de l'IgE-réactivité. Cependant, ces épitopes sont difficiles à localiser *in vitro*, du fait de la difficulté expérimentale d'obtention des complexes allergènes-IgE. C'est pourquoi d'autres approches ont été mises au point, telles que l'introduction de mutations par mutagenèse dirigée dans les zones de surface des allergènes (Ferreira et al., 1998; Chan et al., 2008), l'inhibition de l'interaction entre l'allergène natif et muté avec les IgE à l'aide des anticorps monoclonaux (Chapman et al., 1987; Naik et al., 2008) ou encore le phage display (Folgori et al., 1994; Hantusch et al., 2004). Plusieurs stratégies ont été envisagées pour le traitement de l'allergie (figure n°10). La modification de la structure tridimensionnelle des antigènes a été un moyen efficace pour concevoir des variants hypoallergéniques recombinants qui ont un potentiel d'application clinique et thérapeutique chez les patients allergiques. Les allergènes recombinants structurellement modifiés peuvent être préparés en utilisant des peptides contenant les épitopes T, les épitopes B, ou des fragments d'allergènes dont la structure est altérée soit par mutation des ponts cystéines, par exemple, ou en les fixant sur des molécules porteuses (supports) (Focke et al., 2001; Smith & Chapman, 1996; Takai et al., 1997; Chen et al., 2012a; Chen et al., 2012b; Valenta et al., 2011).



**Figure n°10:** Approches pour la génération de vaccins hypoallergéniques contre les allergies. Au dessus, une représentation schématique de la structure primaire des allergènes. Les peptides impliqués dans la formation d'épitopes B conformationnels sont colorés en orange et rouge, tandis que des peptides reconnus par les lymphocytes T sont en vert, bleu et jaune. (A) Peptides des épitopes T isolés. (B) Fragments d'allergènes recombinants dont les épitopes B sont absents. (C) Oligomères d'allergènes. (D) Mutation des résidus impliqués dans la liaison avec les IgE sans modification de la structure. (E) Variants de repliement. (G) mosaïques contenant des épitopes T conservés tandis que les épitopes B conformationnels sont perdus. (F) Dérivés d'allergènes fixés sur un support avec absence d'épitopes B et des épitopes T (Issue de Valenta *et al.*, 2011).

Au vu de leur grande allergénicité, Der p 1 et Der p 2 ont fait l'objet de plusieurs études pour le développement des hypoallergènes et la conception des vaccins thérapeutiques (Asturias *et al.*, 2009; Zhong *et al.*, 2011; Walgraffe *et al.*, 2009; Chen *et al.*, 2012a; Chen *et al.*, 2012b).

#### I.4 Méthodes Bioinformatiques d'aide à la compréhension des mécanismes liés à l'allergie

Les progrès en biologie moléculaire ont permis un accroissement rapide des données liées à l'allergie (l'augmentation du nombre de séquences des allergènes, leurs structures moléculaires, leurs réactions croisées), ce qui, par conséquent, nécessite des moyens avancés pour leur gestion et leur analyse. Ainsi, des méthodes bioinformatiques attribuées à l'étude des allergènes ont été élaborées, entre autre, la conception des banques et des bases de données

des allergènes. La bioinformatique ou l'immunoinformatique offre également des atouts cruciaux pour la détermination de l'allergénicité des allergènes grâce au développement des méthodes de prédiction de leurs déterminants antigéniques (épitopes B ou T) et des méthodes de prédiction des structures tertiaires et d'amarrage protéines-protéines/protéine-ligands qui aident à la compréhension de la fonction biologique et des mécanismes par lesquels les allergènes pourraient déclencher la réponse immunitaire.

#### I.4.1 Bases de données généralistes et spécialisées

Un grand nombre de séquences, entre autres les séquences des allergènes, sont actuellement disponibles dans la littérature. Ces séquences sont collectées et stockées de manière structurée dans un certain nombre de banques de données. Les principales banques de données sont Genbank, EMBL (*The European Molecular Biology Laboratory*), DDBJ (*DNA Data Bank of Japan*), Uniprot et PDB (*Protein Data Bank*) (Tableau n°3) (Brusic et al., 2003).

**Tableau n° 3:** Bases de données généralistes

Bases de données	URL	Références
Genbank/GenPept	<a href="http://www.ncbi.nlm.nih.gov/genbank/">http://www.ncbi.nlm.nih.gov/genbank/</a>	Benson et al., 2011
EMBL/TrEMBL	<a href="http://www.ebi.ac.uk/embl/">http://www.ebi.ac.uk/embl/</a>	Kulikova et al., 2004; Cochrane et al., 2008
DDBJ/DAD	<a href="http://www.ddbj.nig.ac.jp">www.ddbj.nig.ac.jp</a>	Kaminuma et al., 2011
UniProtKB/Swiss-Prot	<a href="http://www.uniprot.org">http://www.uniprot.org</a>	The UniProt Consortium 2010
PDB	<a href="http://www.rcsb.org">www.rcsb.org</a>	Berman et al., 2000

Les trois banques de données GenBank, EMBL et DDBJ contiennent toutes les séquences nucléiques et leurs annotations (informations sur les séquences, l'organisme producteur, la référence, etc...). Les données sont collectées, mises à jour et échangées quotidiennement entre ces trois banques de données. GenPept, TrEMBL et DAD, trouvés dans ces trois banques de données, respectivement, sont des banques de données de séquences protéiques dérivées à partir des traductions des séquences codantes de ces trois bases de données de nucléotides. Par ailleurs, les séquences protéiques proprement dites sont stockées et annotées dans la banque de donnée Uniprot. Cette dernière fournit des informations sur la fonction des protéines, leurs modifications post-traductionnelles, leurs domaines et leurs sites catalytiques. Les structures de biomolécules sont, quant à elles, déposées dans la banque de donnée PDB.

Toutes ces banques de données disposent des liens entre elles, mais également vers d'autres bases de données rendant ainsi la recherche beaucoup plus enrichie.

Ces bases de données contiennent une proportion importante d'allergènes. En 2012, on compte 896 séquences et 413 structures d'allergènes stockées dans Uniprot et dans la PDB, respectivement. Toutefois, dans le but d'accélérer la recherche dans le domaine de l'allergie, des bases de données spécialisées (quelques unes sont citées dans le tableau n°4) ont été développées pour stocker les informations liées aux allergènes (les données moléculaires, biochimiques et cliniques) (Brusic *et al.*, 2003; Schein *et al.*, 2007). En général, ces bases de données contiennent au moins une liste d'allergènes et des outils bioinformatiques pour la prédiction de leurs épitopes et de l'allergénicité des protéines.

La base de donnée WHO/IUIS, approuvée par l'Organisation Mondiale de la Santé et l'Union Internationale des Associations d'Immunologie, est le site officiel de la nomenclature systématique des allergènes. On y retrouve la liste des allergènes et des isoallergènes. SDAP (*Structural Database of allergen Proteins*) et Allergome sont les bases de données les plus vastes (Ivanciuc *et al.*, 2003; Mari *et al.*, 2006). Elles fournissent toutes les informations disponibles sur les allergènes cités dans la banque de données WHO/IUIS (le nom de l'allergène, la source, la séquence, la structure, les épitopes, les références bibliographiques) et sur les allergènes qui ne sont pas cités dans IUIS mais décrits comme étant des allergènes dans leurs références bibliographiques. La base de donnée AllFam (Radauer *et al.*, 2008) permet la classification des allergènes en les groupant en famille selon leurs propriétés structurelles et fonctionnelles. Elle fusionne les informations trouvées dans la banque de données Allergome (Mari *et al.*, 2006) et la base de données de familles de protéines Pfam (Protein families) (<http://pfam.sanger.ac.uk/>; Punta *et al.*, 2012). Les bases de données AllergenOnline (qui se trouve dans FARRP- *The Food Allergy Research and Resource Program*, Hileman *et al.*, 2002), et ADFS (*Allergen Database for Food Safety*, Nakamura *et al.*, 2009) ont pour but l'amélioration de la sécurité des produits alimentaires pour les consommateurs souffrant d'allergies alimentaires. Les bases de données EVALLER (Barrio *et al.*, 2007), AllerHunter (Muh *et al.*, 2009) et SORTALLER (Zhang *et al.*, 2012a) permettent la prédiction de l'allergénicité des protéines en faisant des comparaisons entre les motifs et les peptides des séquences des allergènes connus. Epitome (*Database of Structurally inferred Antigenic Epitopes in Proteins*, Schlessinger *et al.*, 2006) et CED (*Conformational Epitope Database*, Huang & Honda, 2006) sont des bases de données d'épitopes conformationnels.

Elles fournissent des informations quant à la séquence et la structure de l'antigène, la position des épitopes, la méthode expérimentalement utilisée pour leur identification, leurs interactions avec les anticorps correspondants, et leurs références bibliographiques. Par ailleurs, IEDB (*Epitope Database Analysis Resource*, Vita et al., 2010) et AlgPred (*Prediction of Allergenic Proteins and Mapping of IgE Epitopes*, Saha & Raghava, 2006a) contiennent un ensemble d'outils de prédiction des épitopes B et T ainsi que leurs analyses.

**Tableau n°4:** Bases de données spécialisées en allergies

Base de données	URL	Informations trouvées	Référence
WHO/IUIS	<a href="http://www.allergen.org">http://www.allergen.org</a>	Liste des noms officiels des allergènes	Chapman et al., 2007
SDAP	<a href="http://fermi.utmb.edu/SDAP/">http://fermi.utmb.edu/SDAP/</a>	Séquences d'allergènes, outils de prédiction d'allergénicité et de réactivité croisée des allergènes, collection des épitopes	Ivanciuc et al., 2003
Allergome	<a href="http://www.allergome.org">http://www.allergome.org</a>	Informations sur les allergènes, leur classification, relation entre la structure des allergènes et leurs épitopes	Mari et al., 2006
AllFam	<a href="http://www.meduniwien.ac.at/allergens/allfam/about.html">http://www.meduniwien.ac.at/allergens/allfam/about.html</a>	Classification des allergènes	Radauer et al., 2008
AllergenOnline/ FARRP	<a href="http://www.allergenonline.org/">http://www.allergenonline.org/</a>	Liste des allergènes alimentaires et outils de prédiction d'allergénicité des protéines qui pourraient présenter un risque potentiel de réactivité croisée allergénique	Hileman et al., 2002
ADFS	<a href="http://allergen.nihs.go.jp/ADFS/">http://allergen.nihs.go.jp/ADFS/</a>	Liste des allergènes et isoallergènes alimentaires, liens vers d'autres bases de données, et outils de prédiction d'allergénicité des protéines	Nakamura et al., 2009
EVALLER	<a href="http://www.slv.se/en-gb/Group1/Food-Safety/e-Testing-of-protein-allergenicity/e-Test-allergenicity/">http://www.slv.se/en-gb/Group1/Food-Safety/e-Testing-of-protein-allergenicity/e-Test-allergenicity/</a>	Prédiction de la réactivité croisée en se basant sur la similarité des protéines à des allergènes connus	Barrio et al., 2007
AllerHunter	<a href="http://tiger.dbs.nus.edu.sg/AllerHunter/index.html">http://tiger.dbs.nus.edu.sg/AllerHunter/index.html</a>	Prédiction de la réactivité croisée en se basant sur la similarité des protéines à des allergènes connus	Muh et al., 2009
SORTALLER	<a href="http://sortaller.gzhmc.edu.cn/mail/index.html">http://sortaller.gzhmc.edu.cn/mail/index.html</a>	Prédiction de la réactivité croisée en se basant sur la similarité des protéines à des allergènes connus	Zhang et al., 2012a
IEDB	<a href="http://tools.immuneepitope.org/main/index.html">http://tools.immuneepitope.org/main/index.html</a>	Outils de prédiction des épitopes B et T	Vita et al., 2010
AlgPred	<a href="http://www.imtech.res.in/raghava/algpred/">http://www.imtech.res.in/raghava/algpred/</a>	Outils de prédiction d'allergénicité des protéines et des épitopes sur la base de la séquence	Saha & Raghava, 2006a
Epitome	<a href="http://www.rostlab.org/services/epitome/">http://www.rostlab.org/services/epitome/</a>	Base de données des épitopes conformationnels	Schlessinger et al., 2006
CED	<a href="http://immunet.cn/ced/">http://immunet.cn/ced/</a>	Base de données des épitopes conformationnels	Huang & Honda, 2006
MimoDB	<a href="http://immunet.cn/mimodb/">http://immunet.cn/mimodb/</a>	Base de données de mimotopes	Huang et al., 2012

MimoDB, quant à elle, est une base de données de mimotopes. Ces derniers sont des peptides ou des protéines qui imitent les épitopes conformationnels d'un allergène naturel mais qui ne présentent aucune similitude de séquence avec une quelconque partie de ce dernier (Szalai et al., 2008). Ces mimotopes sont sélectionnés à partir des banques de peptides phages (phage display) et sont surtout utilisés dans l'immunothérapie épitope-spécifique.

#### I.4.2 Méthodes de prédiction des épitopes B

Les épitopes sont des régions se trouvant à la surface des allergènes. On en distingue deux types: les épitopes B qui sont reconnus par les anticorps, et les épitopes T qui font réagir les lymphocytes T. Dans les maladies allergiques, les travaux sont surtout consacrés à la détermination des épitopes B des allergènes en raison de leurs interactions avec les IgE spécifiques. Ces épitopes sont soit linéaires, soit conformationnels. Un épitope linéaire (aussi appelé séquentiel ou continu) est un fragment de la séquence constitué d'une suite contigüe d'acides aminés allant de 9 à 17 acides aminés, alors qu'un épitope conformationnel (également appelé discontinu) est un épitope composé de plusieurs fragments dispersés le long de la séquence et rassemblés dans une proximité spatiale lorsque la protéine est repliée sous sa structure tertiaire ou quaternaire (Van Regenmortel, 2001).

Au cours de ces vingt dernières années, les méthodes bioinformatiques ont essentiellement porté sur la prédiction des épitopes linéaires. La plupart de ces méthodes sont basées sur les propriétés physico-chimiques des allergènes telles que la flexibilité (Karplus & Schultz, 1985; Vihinen et al., 1994), l'accessibilité au solvant (Alix, 1999), l'hydrophilie (Parker et al., 1986; Efremov & Alix 1993), la composition en acides aminés (Hopp & Woods, 1981; Hopp, 1993) et la position des tournants (Pellequer et al., 1993). Plusieurs serveurs de prédiction des épitopes linéaires sont accessibles sur le web (tableau n°5).

Les programmes PEOPLE (*Predictive Estimation of Protein Linear Epitopes*, Alix, 1999), BEPITOPE (Odorico & Pellequer, 2003) et Bcepred (Saha & Raghava, 2004) utilisent une combinaison des paramètres physico-chimiques pour la prédiction des épitopes sur la base de la séquence des allergènes. Saha & Raghava (2004) ont montré que les programmes qui se basent sur la combinaison de ces paramètres physico-chimiques ont de meilleures performances de prédiction. Cela dit, l'étude réalisée par Blythe & Flower (2005) sur la corrélation entre les échelles de propensions de ces paramètres physico-chimiques et la localisation des épitopes linéaires dans 50 protéines a montré que ces prédictions sont à peine meilleures que celles du hasard. D'autres méthodes de prédiction, tels ABCpred (Saha &

Raghava, 2006a), Bepipred (Larsen et al., 2006) et BCPred (El-Manzalawy et al., 2008), ont été améliorées en utilisant des algorithmes d'apprentissage automatisé comme les machines à vecteur de support ou les algorithmes génétiques. Malheureusement, les performances de prédiction de ces méthodes ne sont pas assez satisfaisantes (Liang et al., 2009). En outre, plus de 90% des épitopes sont conformationnels (Barlow et al., 1986; Chen et al., 2011). Les bases de données IEDB et CED (tableau n°4) collectent les structures des complexes antigène-anticorps stockées dans la PDB. Ces complexes sont exploités dans les méthodes de prédiction des épitopes conformationnels. Cependant, vu le faible nombre des structures des allergènes résolues expérimentalement, un petit nombre d'outils tenant compte des propriétés structurales ont été développés jusqu'à présent pour la prédiction des épitopes conformationnels (Liu & Hu 2011) (tableau n°5).

CEP (*Conformational epitope prediction*, Kulkarni-Kale et al., 2005) et PEPOP (Moreau et al., 2008) utilisent l'accessibilité au solvant des résidus pour prédire les épitopes conformationnels. DiscoTope (Andersen et al., 2006) et PEPITO (Sweredoski & Baldi, 2008) se basent sur la distribution spatiale et l'accessibilité au solvant des résidus en surface. Par ailleurs, plus d'éléments conformationnels, comme la compacité spatiale des résidus en surface, sont pris en compte dans le programme SEPPA (*Spatial Epitope Prediction of Protein Antigens*, Sun et al., 2009) et EPCES (*Epitopes on Protein Surfaces by Consensus Scoring*, Liang et al., 2009, Liang et al., 2009). Dall'Antonia et al. (2011) ont développé le programme SPADE (*Surface comparison-based Prediction of Allergenic Discontinuous Epitopes*) pour la prédiction des épitopes B conformationnels en se basant sur les propriétés structurales et immunologiques des allergènes en prenant en compte leur réactivité croisée. Toutefois, l'application des méthodes basées sur la structure est limitée par le faible nombre des structures des allergènes disponibles. Ainsi, des approches de prédiction des épitopes conformationnels sur la base de la structure primaire des allergènes ont été développées. De cette façon, les programmes CBTOPE (Ansari & Raghava 2010) et Bcell-prediction (Zhang et al., 2012b) ont été récemment élaborés en se basant sur des dérivés potentiels des caractéristiques de séquences (propriétés physico-chimiques, profil d'évolution, de composition, de nature et d'appariement des acides aminés) et d'algorithmes d'apprentissage pour la prédiction des épitopes conformationnels (Ansari & Raghava, 2010; Zhang et al., 2012b). En outre, d'autres programmes comme Epitopia (Rubinstein et al., 2009) et Ellipro (Ponomarenko et al., 2008) prédisent les épitopes sur la base de la structure ou de leurs séquences des allergènes.

Tableau n°5: Outils de prédictions des épitopes B

Programme	URL	Description	Référence
PEOPLE	/	Prédiction des épitopes B linéaires, basée sur les échelles de propension des propriétés physico-chimiques des acides aminés	Alix, 1999
BEPITOPE	/	Prédiction des épitopes B linéaires, basée sur les échelles de propension des propriétés physico-chimiques des acides aminés	Odorico & Pellequer, 2003
Bcepred	<a href="http://www.imtech.res.in/raghava/bcepred">http://www.imtech.res.in/raghava/bcepred</a>	Prédiction des épitopes B linéaires, basée sur les échelles de propension des propriétés physico-chimiques des acides aminés	Saha & Raghava, 2004
ABCpred	<a href="http://www.imtech.res.in/raghava/abcpred">http://www.imtech.res.in/raghava/abcpred</a>	Prédiction des épitopes B linéaires, basée sur un algorithme d'apprentissage (réseau de neurones)	Saha & Raghava, 2006b
Bepipred	<a href="http://www.cbs.dtu.dk/services/BepiPred">http://www.cbs.dtu.dk/services/BepiPred</a>	Prédiction des épitopes B linéaires, basée sur un algorithme d'apprentissage (combinaison entre le modèle de Markov caché et une échelle de propension de l'hydrophilicité)	Larsen et al., 2006
BCPred	<a href="http://ailab.cs.iastate.edu/bcpred/">http://ailab.cs.iastate.edu/bcpred/</a>	Prédiction des épitopes B linéaires, basée sur un algorithme d'apprentissage (machines à vecteur de support)	El-Manzalawy et al., 2008
CEP	<a href="http://bioinfo.ernet.in/cep.htm">http://bioinfo.ernet.in/cep.htm</a>	Prédiction des épitopes B linéaires et conformationnels sur la base de la structure 3D des allergènes	Kulkarni-Kale et al., 2005
PEPOP	<a href="http://www.webcitation.org/quey.php?url=http://diagtools.sysdiag.cnrs.fr/PEPOP/&amp;refdoi=10.1186/1471-2105-9-71">http://www.webcitation.org/quey.php?url=http://diagtools.sysdiag.cnrs.fr/PEPOP/&amp;refdoi=10.1186/1471-2105-9-71</a>	Prédiction des épitopes B conformationnels, sur la base de la structure 3D des allergènes	Moreau et al., 2008
DiscoTope	<a href="http://www.cbs.dtu.dk/services/DiscoTope">http://www.cbs.dtu.dk/services/DiscoTope</a>	Prédiction des épitopes B conformationnels, sur la base de la structure 3D des allergènes	Andersen et al., 2006
PEPITO	<a href="http://pepito.proteomics.ics.uci.edu">http://pepito.proteomics.ics.uci.edu</a>	Prédiction des épitopes B conformationnels sur la base de la structure 3D des allergènes	Sweredoski & Baldi, 2008
SEPPA	<a href="http://lifecenter.sgst.cn/seppa/">http://lifecenter.sgst.cn/seppa/</a>	Prédiction des épitopes B conformationnels sur la base de la structure 3D des allergènes	Sun et al., 2009
EPCES	<a href="http://www.t38.physik.tu-muenchen.de/programs.htm">http://www.t38.physik.tu-muenchen.de/programs.htm</a>	Prédiction des épitopes B conformationnels sur la base de la structure 3D des allergènes	Liang et al., 2009
SPADE	/	Prédiction des épitopes B conformationnels sur la base de la structure 3D et réactivité croisée des allergènes	Dall'Antonia et al., 2011
CBTOPE	<a href="http://www.imtech.res.in/raghava/cbtope/">http://www.imtech.res.in/raghava/cbtope/</a>	Prédiction des épitopes B conformationnels sur base de la séquence en utilisant un algorithme d'apprentissage (machines à vecteur de support)	Ansari & Raghava 2010
Bcell-prediction	<a href="http://bcell.whu.edu.cn">http://bcell.whu.edu.cn</a>	Prédiction des épitopes B conformationnels sur la base de la séquence des allergènes	Zhang et al., 2012b
Epitopia	<a href="http://epitopia.tau.ac.il/">http://epitopia.tau.ac.il/</a>	Prédiction des épitopes B sur la base de la structure ou de la séquences en utilisant un algorithme d'apprentissage (machines à vecteur de support)	Rubinstein et al., 2009
EllipPro	<a href="http://tools.immuneepitope.org/tools/ElliPro/iedb_input">http://tools.immuneepitope.org/tools/ElliPro/iedb_input</a>	Prédiction des épitopes B sur la base de la structure 3D des allergènes ou de leurs séquences	Ponomarenko et al., 2008
EPSVR/EPMeta	<a href="http://sysbio.unl.edu/EPSVR">http://sysbio.unl.edu/EPSVR</a> <a href="http://sysbio.unl.edu/EPMeta">http://sysbio.unl.edu/EPMeta</a>	Prédiction des épitopes B conformationnels en utilisant un algorithme d'apprentissage (la régression par machines à vecteurs de support)	Liang et al., 2010
PepSurf/Mapitope	<a href="http://pepitope.tau.ac.il/">http://pepitope.tau.ac.il/</a>	Prédiction des épitopes B conformationnels (Phage-display)	Mayrose et al., 2007
EPIMAP	/	Prédiction des épitopes B conformationnels (Phage-display)	Mumey et al., 2003 / Mumey et al., 2006
MimoPro	<a href="http://informatics.nenu.edu.cn/MimoPro/">http://informatics.nenu.edu.cn/MimoPro/</a>	Prédiction des épitopes B conformationnels (Phage-display)	Chen et al., 2011
Per-3D-Search	<a href="http://kyc.nenu.edu.cn/Pep3DSearch/">http://kyc.nenu.edu.cn/Pep3DSearch/</a>	Prédiction des épitope B conformationnels sur la base de l'analyse des mimotopes (Phage-display)	Huang et al., 2008

Epitopia s'appuie sur l'ensemble des paramètres physico-chimiques et structuraux qui caractérisent un épitope, tels que la préférence en acides aminés, le type de structure secondaire, la forme géométrique, et la conservation des résidus au cours de l'évolution, décrits par Rubinstein *et al.* (2008). Le programme Ellipro, quant à lui, représente la surface des protéines par des patchs en ellipsoïde et calcule par la suite un index lié au pourcentage des atomes enfermés dans ces patchs pour chaque résidus en surface. Ce programme prédit les épitopes des allergènes en se basant sur leurs structures 3D ou sur la base de leurs séquences si ces protéines présentent une homologie avec d'autres allergènes ayant des structures connues.

Liang *et al.* (2010) ont développé le programme EPSVR qui utilise des machines à vecteurs de support pour la prédiction des épitopes conformationnels. Ces auteurs ont également conçu un meta serveur (EPMeta) qui regroupe les résultats de prédiction de plusieurs programmes existants (EPSVR, EPCES, EPITOPIA, SEPPA, PEPITO et Discotope).

D'autre part, une autre approche qui combine à la fois les techniques expérimentales et informatiques pour l'identification des épitopes B a été également proposée (Pizzi *et al.*, 1995; Smith & Petrenko, 1997). Dans cette méthode une bibliothèque de phage-display de peptides aléatoires est scannée contre un anticorps d'intérêt afin d'obtenir un panel de peptides (appelés mimotopes) qui se lie à l'anticorps avec une affinité élevée (Meloan *et al.*, 2000; El-Manzalawy & Honavar, 2010). Les programmes d'analyse des mimotopes disponibles dans la littérature tels PepSurf et Mapitope (Mayrose *et al.*, 2007), EPIMAP (Mumey *et al.*, 2006), MimoPro (Chen *et al.*, 2011) et Per-3D-Search (Huang *et al.*, 2008) diffèrent les uns des autres en termes de la façon dont ils représentent la structure ou la séquence de l'antigène, de la façon dont ils alignent les mimotopes avec la structure ou la séquence de l'antigène cible et de la façon dont ils groupent les mimotopes et classent les épitopes prédits.

Etant donné que la majorité des épitopes B sont conformationnels, les méthodes de prédictions les plus prometteuses seront celles qui se fondent sur les bases de données qui contiennent le maximum de complexes épitope-paratope (antigène-anticorps) (Ponomarenko & Bourne; 2007; Liang *et al.*, 2009). L'étude de la structure 3D conduit donc à une amélioration significative de la prédiction des épitopes.

### I.4.3 Prédiction de la structure des allergènes

La structure 3D est très importante puisqu'elle détermine les propriétés biochimiques et immunologiques des allergènes. En général, les structures 3D des protéines sont déterminées soit par cristallographie aux rayons-X soit par RMN. Cependant, ces méthodes sont coûteuses et restent un processus long en terme de temps. En complément à ces approches expérimentales, les méthodes de modélisation moléculaire pour la prédiction des structures tertiaires ont été développées. Plusieurs études ont montré l'importance de ces méthodes de modélisation dans la prédiction de la structure, de l'allergénicité et de la réactivité croisée entre les allergènes (Roy *et al.*, 2003; Barre *et al.*, 2005; Chapman *et al.*, 2007; Oezguen *et al.*, 2009; Pomès, 2010; Ghosh & Gupta-Bhattacharya, 2012). Shen *et al.*, (2011) ont prédit la structure et ont pu identifier les épitopes de l'allergène Der f 7 en utilisant des méthodes bioinformatiques d'alignement de séquences et de modélisation moléculaire.

Trois types de méthodes sont appliqués pour la prédiction de la structure tertiaire des protéines: la modélisation comparative, la reconnaissance de repliement (*threading*) et les méthodes *ab initio*.

La modélisation comparative est fondée sur le fait que deux protéines qui partagent une identité de séquence élevée (plus de 30%) adoptent un même repliement (Rost, 1999; Baker et Sali, 2001). Son principe consiste à aligner la séquence d'un allergène donné dont la structure est inconnue (cible) avec la séquence d'une ou de plusieurs protéines ayant une structure expérimentale (RMN ou Rayons-X) (patrons ou références) (Sali & Blundell, 1993). Les programmes de prédiction des structures des protéines par modélisation comparative élaborés jusqu'à ce jour se basent sur différentes approches: assemblage de corps rigides par SWISS-MODEL (<http://swissmodel.expasy.org/> ; Schwed *et al.*, 2003), correspondance de segments par SEGMOD (Levitt, 1992), ou satisfaction des contraintes spatiales par MODELLER (<http://salilab.org/modeller/modeller.html> ; Sali et Blundell, 1993).

La méthode de reconnaissance de repliement est préconisée quand la similarité de séquence entre la cible et les patrons est comprise entre 20 et 30%. Elle consiste à enfileur la séquence de la cible sur une bibliothèque de repliement afin de déterminer les structures qui correspondent le mieux à la séquence sur la base d'un critère énergétique ou de score. Les programmes de reconnaissance de repliement les plus populaires sont 3D-PSSM ou Phyre

(<http://www.sbg.bio.ic.ac.uk/~3dpssm/>; Kelley et al., 1999; Kelley & Sternberg, 2009) et THREADER ([http://bioinf.cs.ucl.ac.uk/software\\_downloads/threader/](http://bioinf.cs.ucl.ac.uk/software_downloads/threader/); Jones, 1998).

Par ailleurs, si peu de similarité de séquence est détectée, les méthodes *ab initio* sont utilisées pour prédire la structure de la cible à partir de sa séquence en acides aminés en se basant sur leurs interactions physicochimiques (électrostatiques, van der Waals, liaisons hydrogènes, etc...) entre les atomes des résidus (Cornell et al., 1995). Ces méthodes se basent sur l'hypothèse que la structure native des protéines correspond à l'énergie libre minimale. La méthode de prédiction *ab initio* ROSETTA est la plus performante (Rohl et al. 2004; Jauch et al., 2007).

Les performances de ces méthodes sont évaluées régulièrement par le concours CASP (*Critical Assessment of techniques for protein Structure Prediction*) (<http://www.predictioncenter.org/casp10/index.cgi> ; Moult et al., 2011). Malgré les progrès réalisés dans chacune de ces méthodes de prédiction, la modélisation comparative demeure la plus performante (Oezguen et al., 2009; Shen et al., 2011; Ghosh & Gupta-Bhattacharya, 2011).

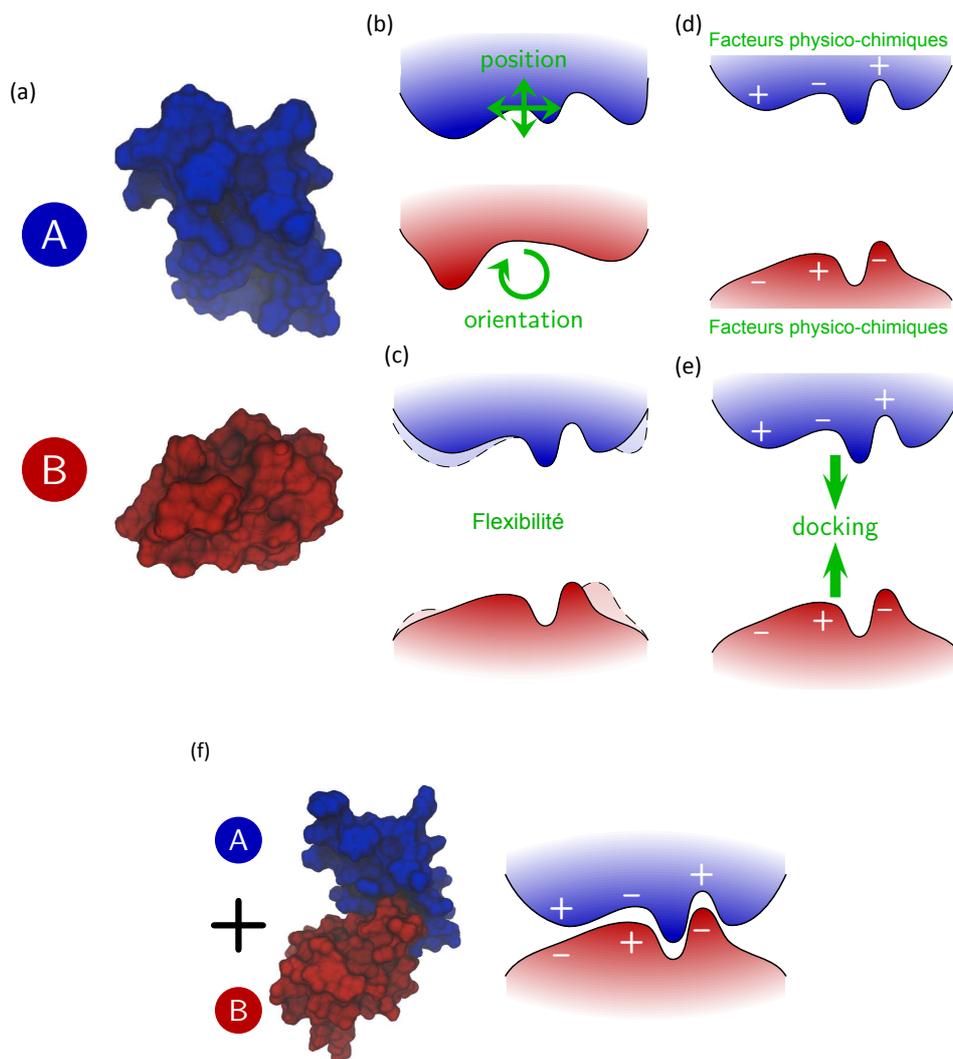
D'autre part, plusieurs programmes ont été développés pour évaluer la qualité et raffiner les modèles construits par ces méthodes de prédiction. PROCHECK (<http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/>; Laskowski, 1993) évalue la qualité géométrique (stéréochimique) des modèles. Anolea (Melo & Feytmans, 1998) et DOPE (Shen & Sali, 2006) se basent sur des potentiels statistiques et Eval123D est fondé sur des matrices de compatibilité 1D/3D ([http://bioserv.cbs.cnrs.fr/HTML\\_BIO/frame\\_valid.html](http://bioserv.cbs.cnrs.fr/HTML_BIO/frame_valid.html); Gracy et al., 1993).

#### **I.4.4 Amarrage protéine-protéine/protéine-ligands**

La localisation des sites de liaisons ou d'interactions entre les protéines ou entre les protéines et des ligands est d'une importance fondamentale pour la compréhension de la fonction biochimique des allergènes, la détermination de leurs épitopes conformationnels et la conception des médicaments (Joseph-McCarthy, 1999; Zhou & Qin 2007). En effet, plusieurs études ont montré le rôle de l'oligomérisation des allergènes dans leur allergénicité soit par leur susceptibilité de créer le pontage aux IgE à la surface des mastocytes (comme c'est le cas pour l'allergène majeur du pollen de bouleau Bet v 1) soit par la création des cavités qui

pourraient accueillir et transporter des ligands (comme les allergènes d'acariens Der p 2 et Der f 2) (Lascombe et *al.*, 2000; Schöll et *al.* 2005; de Halleux et *al.*, 2006; Kaiser et *al.*, 2007; Chapman, 2008; Mueller et *al.*, 2010a; Mueller et *al.*, 2010b; Pomés, 2010).

L'amarrage de protéines (appelé aussi le *docking* ou l'ancrage) consiste à prédire la conformation optimale stable d'un complexe de protéines ou protéine-ligand à partir des structures individuelles des partenaires dans leur état non lié. La stabilité de ces complexes est essentiellement évaluée selon la complémentarité géométrique (ou structurelle) et la complémentarité chimique (Figure n°11) (Campbell et *al.*, 2003; Zhou & Qin, 2007; Ezkurdia et *al.*, 2009). Le principe de la complémentarité géométrique est l'assemblage des molécules afin de trouver les parties qui s'ajustent le mieux les unes avec les autres (Figure n°11 a, b et c). La complémentarité chimique, quant à elle, prend en compte les interactions chimiques entre les molécules à amarrer tels que les ponts hydrogènes, les interactions électrostatiques, les contacts hydrophobes et le phénomène de solvation/désolvation (Figure n°11d). La combinaison de ces paramètres conduit à la formation du complexe moléculaire (Figure n°11 e et f) (Simard, 2012).



**Figure n°11:** Mesures pour l'évaluation de l'amarrage protéine "A" - protéine "B" (Issue de Simard, 2012). (a) Protéines A et B avant l'amarrage. (b) et (c) Prise en compte de la complémentarité géométrique pour l'amarrage. (d) et (e) Prise en compte de la complémentarité chimique pour l'amarrage. (f) Résultat de l'amarrage des deux molécules A et B.

Néanmoins, le nombre de combinaison de ces critères géométriques et chimiques fait de l'amarrage moléculaire une tâche très complexe et très coûteuse en terme de calcul et de temps. Pour répondre à ces difficultés, plusieurs programmes ont été élaborés en utilisant différentes approches pour tenter de trouver des solutions d'amarrage protéine-protéine/protéine-ligand optimales. Ces programmes diffèrent entre eux selon (i) la fonction de score qui permet d'évaluer l'affinité entre les molécules amarrées en calculant leur énergie d'interaction, (ii) l'algorithme d'exploration conformationnelle et (iii) la prise en compte de la

flexibilité des molécules. Selon cette dernière, on distingue trois types d'amarrage moléculaire: (1) l'amarrage rigide, où aucune molécule n'est considérée flexible, (2) l'amarrage semi-flexible où une seule molécule est considérée flexible, et (3) l'amarrage flexible où les deux molécules sont considérées flexibles. Dans le tableau n°6 nous avons listé les programmes les plus populaires d'amarrage protéine-protéine (PatchDock, 3D-Garden, ClusPro, SmoothDock, GRAMM-X, ZDOCK, HADDOCK et Rosetta) et protéine-ligand (AutoDock, EADock, GOLD et FlexX).

Les deux programmes d'amarrage PatchDock (Schneidman-Duhovny et *al.*, 2005) et 3D-Garden (Lesk & Sternberg 2008) sont basés sur la complémentarité géométrique des protéines considérées comme corps rigides. ClusPro (Comeau et *al.*, 2004), SmoothDock (Camacho et Champ, 2006) et GRAMM-X (Tovchigrechko & Vakser, 2006) sont également des programmes d'amarrage corps rigides mais ils s'appuient sur des techniques de corrélation de la transformée de Fourier rapide (ou "*Fast Fourier Transform*," FFT) qui permettent de calculer rapidement les complémentarités géométrique et chimique des protéines dans une grille 3D cartésienne en faisant des mouvements de translations/rotations des molécules. ZDOCK (Chen et *al.*, 2003; Pierce et *al.*, 2011), quant à lui, applique ces techniques mais en prenant en considération la flexibilité des deux protéines à docker. Par ailleurs, HADDOCK (*High Ambiguity Driven protein-protein DOCKing*, Dominguez et *al.*, 2003) et Rosetta (Gray et *al.*, 2003) font appel à la dynamique moléculaire pour modéliser la flexibilité des protéines.

AutoDock (Morris et *al.*, 1998), EADock (*Evolutionary Algorithm for Docking*, Grosdidier et *al.*, 2007) et GOLD (Jones et *al.*, 1997) sont des programmes d'amarrage protéine-ligand qui tiennent en compte la flexibilité de ces molécules. Ils utilisent un algorithme génétique ou évolutionnaire pour effectuer la recherche dans l'espace conformationnel des conformations qui sembleraient être de plus basse énergie libre. Par ailleurs, FlexX (Kramer et *al.*, 1999) emploie la technique de la construction incrémentale (construction à partir des fragments) pour l'amarrage du ligand au sein du site actif de la protéine.

Le concours CAPRI (*Critical Assessment of PRediction of Interactions*) (<http://www.ebi.ac.uk/msd-srv/capri/>) permet l'évaluation de la performance de ces méthodes d'amarrage.

**Tableau n° 6:** Liste non-exhaustive des programmes d'amarrage protéine-protéine/protéine-ligands les plus populaires

Programme	URL	Description de la méthode	Référence
PatchDock	<a href="http://bioinfo3d.cs.tau.ac.il/PatchDock/">http://bioinfo3d.cs.tau.ac.il/PatchDock/</a>	Amarrage protéine-protéine, corps rigides, basée sur la complémentarité géométrique	Schneidman-Duhovny <i>et al.</i> , 2005
3D-Garden	<a href="http://www.sbg.bio.ic.ac.uk/~3d-garden/">http://www.sbg.bio.ic.ac.uk/~3d-garden/</a>	Amarrage protéine-protéine, corps rigides, basée sur la complémentarité géométrique	Lesk & Sternberg 2008
ClusPro	<a href="http://cluspro.bu.edu">http://cluspro.bu.edu</a>	Amarrage protéine-protéine, corps rigides, basée sur la transformée de Fourier rapide (FFT)	Comeau <i>et al.</i> , 2004
SmoothDock	<a href="http://structure.pitt.edu/servers/smoothdock/">http://structure.pitt.edu/servers/smoothdock/</a>	Amarrage protéine-protéine, corps rigides, basée sur FFT	Camacho et Champ, 2006
GRAMM-X	<a href="http://vakser.bioinformatics.ku.edu/resources/gramm/grammx/">http://vakser.bioinformatics.ku.edu/resources/gramm/grammx/</a>	Amarrage protéine-protéine, corps rigides, basé sur FFT	Tovchigrechko & Vakser, 2006
ZDOCK	<a href="http://zdock.umassmed.edu">http://zdock.umassmed.edu</a>	Amarrage protéine-protéine, flexibles, basée sur FFT	Chen <i>et al.</i> , 2003/ Pierce <i>et al.</i> , 2011
RosettaDock	<a href="http://c4c.uwc4c.com/express_license_technologies/rosetta">http://c4c.uwc4c.com/express_license_technologies/rosetta</a>	Amarrage protéine-protéine, corps flexibles	Gray <i>et al.</i> , 2003
HADDOCK	<a href="http://haddock.chem.uu.nl">http://haddock.chem.uu.nl</a>	Amarrage protéine-protéine, corps flexibles	Dominguez <i>et al.</i> , 2003
AutoDock	/	Protéines-ligands flexibles, basée sur l'algorithme génétique	Morris <i>et al.</i> , 1998
EADock	/	Protéines-ligands, basée sur l'algorithme génétique	Grosdidier <i>et al.</i> , 2007
GOLD	<a href="http://www.ccdc.cam.ac.uk/products/life_sciences/gold/">http://www.ccdc.cam.ac.uk/products/life_sciences/gold/</a>	Amarrage protéines-ligands, flexibles, basée sur l'algorithme génétique	Jones <i>et al.</i> , 1997
FlexX	<a href="http://www.biosolveit.de/FlexX/">http://www.biosolveit.de/FlexX/</a>	Amarrage protéines (rigide)-ligands (flexibles), basée sur construction incrémentale	Kramer <i>et al.</i> , 1999

La contribution des méthodes bioinformatiques dans la connaissance des états oligomériques des allergènes, le criblage ou le screening *in silico* de petits ligands ainsi que la prédiction de leurs épitopes permettent la compréhension des propriétés des allergènes et offrent de grands avantages dans la conception de vaccins d'hypoallergènes et les traitements d'immunothérapie (Vrtala *et al.*, 2001; Reese *et al.*, 2007; Kundu & Roy, 2010).

## *Chapitre II:*

*Etude de la réactivité croisée  
entre les allergènes des familles 5  
et 21 et prédiction des épitopes de  
Der p 5*

## *Chapitre II:*

### *Introduction et problématique*

## II. 1 Introduction et problématique

L'allergie aux acariens, essentiellement due à *D. pteronyssinus*, *D. farinae* et de *B. tropicalis* est la cause principale des maladies des voies respiratoires et d'eczéma (Chew et al., 1999; Chua et al., 2007; Shek et al., 2010; Thomas, 2011; Jeong et al., 2012). Jusqu'à présent, plus de 20 groupes d'allergènes d'acariens ont été identifiés.

Dans cette étude, nous nous sommes focalisés sur les allergènes d'acariens de poussière de maisons de la famille 5. Plusieurs travaux ont montré la forte allergénicité de l'allergène "Blo t 5" de *Blomia tropicalis* dans les zones tropicales et subtropicales, et l'importance des allergènes Der p 5 et Der f 5 de *Dermatophagoides pteronyssinus* et *Dermatophagoides farinae*, respectivement, dans les maladies respiratoires (Lynch et al., 1997; Thomas et al., 2002; Thomas et al., 2007; Kidon et al., 2011; Cui et al., 2012). D'autre part, Gao et al. (2007) ont isolé un paralogue (produit de copie de gène) de Blo t 5 appelé "Blo t 21". Les allergènes de cette famille 21 (Blo t 21, Der p 21 et Der f 21) connus jusqu'à présent sont donc considérés comme des allergènes homologues à ceux de la famille 5. D'autres allergènes de la famille 5 appartenant à d'autres espèces d'acariens ont également été identifiés: Lep d 5 de *Lepidoglyphus destructor*, Ale o 5 de *Aleuroglyphus ovatus*, Tyr p 5 de *Tyrophagus putrescentiae*, Sui m 5 de *Suidasia medanensis* et Gly d 5 de *Glycyphagus domesticus* (Eriksson et al., 2001; WHO/IUIS/GenBank; Gao et al., 2007), mais peu d'études, voir aucune, n'a été consacrée à l'étude de la prévalence et de la réactivité croisée entre ces derniers (Ale o 5, Tyr p 5, Sui m 5 et Gly d 5), contrairement aux allergènes isolés de *B. tropicalis*, *D. pteronyssinus* et de *D. farinae* (Blo t 5, Blo t 21, Der p 5, Der p 21, Der f 5 et Der f 21).

Les allergènes Blo t 5 et Blo t 21 sont prédominants dans les zones tropicales et subtropicales, alors que Der p 5, Der p 21, Der f 5 et Der f 21 sont surtout trouvés dans les zones tempérées (Shew et al., 1999). La prévalence de Blo t 5 est la plus élevée et son allergénicité est du même ordre que celle des allergènes majeurs Der p 1 et Der p 2 (Kuo et al., 1999; Baqueiro et al., 2006; Chua et al., 2007; Chew et al., 2009; Kidon et al., 2011). Par ailleurs, plusieurs études ont montré également le rôle déterminant de l'allergène Der p 5 dans les allergies respiratoires (Lynch et al., 1998; Kauffman et al., 2006; Thomas et al., 2007; Weghofer et al., 2008b; Thomas, 2011; Thomas, 2012) et dans le développement de l'eczéma chez les enfants (Shek et al., 2010). Cet allergène présente une IgE-réactivité entre 9-70% selon les patients allergiques et les régions où ils se trouvent (surtout en Europe et en Amérique).

Certains travaux ont étudié la réactivité croisée entre les allergènes de la famille 5, entre ceux de la famille 21 et entre les allergènes des ces deux familles (tableau n°7). Le travail de Caraballo et *al.* (1998) est le seul qui a montré l'existence d'une forte réactivité croisée entre Der p 5 et Blo t 5, contrairement à d'autres travaux (Arruda et *al.*, 1997; Chew et *al.*, 1999; Kuo et *al.*, 2003; Simpson et *al.*, 2003) où la réactivité croisée entre ces deux allergènes est estimée de faible à modérée, suggérant ainsi que les allergènes du groupe 5 de *D. pteronyssinus* et de *B. tropicalis* sont spécifiques à l'espèce.

**Tableau n° 7:** Réactivité croisée entre les allergènes des familles 5 et 21

	Der p 21	Der f 21	Blo t 21	Blo t 5	Der p 5	Der f 5	Lep d 5	Tyr p 5	Sui m 5	Ale o 5	Gly d 5
Derp21	/										
Derf21		/									
Blot21			/								
Blot5				/							
Derp5					/						
Derf5						/					
Lepd5							/				
Tyrp5								/			
Suim5									/		
Aleo5										/	
Gly d 5											/

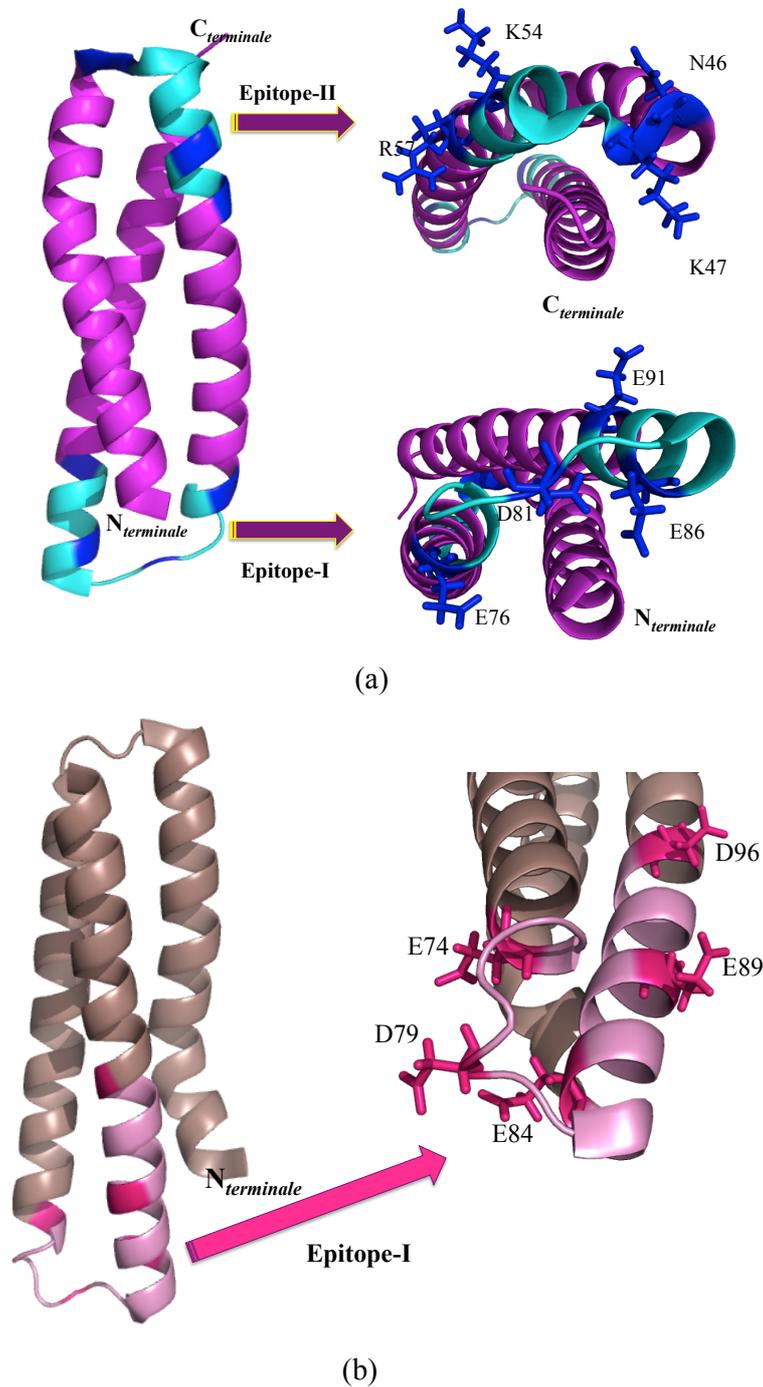
 Faible réactivité croisée     
  Réactivité croisée importante     
  Réactivité croisée non déterminée

Une faible réactivité croisée est également observée entre les deux allergènes Der p 5 et Blo t 5 et entre leurs homologues Der p 21 et Blo t 21 (Weghofer et *al.*, 2008b; Weghofer et *al.* 2008c; Gao et *al.*, 2007; Tan et *al.*, 2012). Weghofer et *al.* (2008b et 2008c) ont montré la faible réactivité croisée entre Der p 5/Blo t 5, Der p 5/ Lep d 5, Der p 21/Blo t 5 et Der p 21/Lep d 5. Le même résultat a été également trouvé par Gao et *al.* (2007) entre Der p 5/Blo t 5, Blo t 5/ Lep d 5 et Der p 5/ Lep d 5. Par ailleurs, Reginald et *al.* (2004) ont mis en évidence une importante réactivité croisée entre Sui m 5 et Blo t 5 et entre Sui m 5 et Der f 5. Cette réactivité croisée entre les allergènes est due à la conservation ou non des épitopes qui se trouvent à la surface de ses allergènes (King & Spaugfort, 2000; Pomés, 2010).

Parmi les allergènes des familles 5 et 21, seuls les épitopes de Blo t 5 et de Blo t 21 ont été caractérisés expérimentalement en utilisant des anticorps monoclonaux (Acm) ou par

mutagénèse dirigée (Chan et *al.*, 2008; Naik et *al.*, 2008; Tan et *al.*, 2012). Les principaux résidus de ces épitopes qui jouent un rôle crucial dans la liaison avec les anticorps ont été également identifiés (figure n°12). Le premier épitope de Blo t 5 (code PDB 2JMH; Naik et *al.*, 2008) est composé de deux résidus majeurs (N46 et K47) situés à l'extrémité de la deuxième hélice et de deux autres résidus majeurs (K54 et R57) dans le début de la troisième hélice (figure n°12a). Le deuxième épitope, quant à lui, comporte quatre résidus majeurs: un dans la fin de la deuxième hélice (E76), un dans la deuxième boucle (D81) et deux dans le début de l'hélice *C\_terminale* (E86 et E91) (Chan et *al.*, 2008) (figure n°12a). L'épitope de Blo t 21 (code PDB 2LM9) se situe dans la même région que le deuxième épitope de Blo t 5. Il est constitué de cinq résidus majeurs: un résidu dans la fin de la deuxième hélice (E74), un résidu dans la deuxième boucle (D79) et trois résidus dans le début de l'hélice *C\_terminale* (E84, E89 et D96) (figure n°12b). Cet épitope de Blo t 21 a été rapporté comme étant un épitope conformationnel contrairement à celui de Blo t 5 qui est décrit comme un épitope linéaire. Cependant, El-Manzalawy & Honavar (2010) ont décrit que même dans le cas des épitopes B linéaires, les interactions anticorps-antigène sont souvent dépendantes de la structure. Un épitope linéaire est en fait défini comme étant un épitope qui peut fixer les IgE même après dénaturation de l'antigène, contrairement à l'épitope conformationnel qui est lié à la structure 3D.

Par ailleurs, plusieurs travaux ont déjà mentionné l'importance de l'allergénicité de la troisième hélice de Blo t 5 (Caraballo et *al.*, 1996; Caraballo et *al.*, 1998; Tsai et *al.*, 2003). De plus, Chan et *al.* (2008) ont montré que la mutation des résidus se trouvant dans l'épitope identifié par Naik et *al.* (2008), ne réduit pas significativement le taux de reconnaissance des IgEs à cet allergène. Cela appuie l'idée de l'importance de l'épitope de Blo t 5 situé au niveau de son hélice *C\_terminale*. Ainsi, selon l'importance de ces épitopes, nous avons nommé l'épitope qui se trouve dans la deuxième boucle-hélice *C\_terminale* "épitope-I", et l'épitope situé dans la première boucle-deuxième hélice "épitope-II" (figure n°12).



**Figure n°12:** Localisation des épitopes de Blo t 5 et de Blo t 21 déterminés expérimentalement. (a) Position des épitopes I et II (en cyan) et de leurs résidus majeurs (sticks bleus) de Blo t 5 (code PDB 2JMH en magenta). (b) Position de l'épitope-I (en rose) et ses résidus majeurs (sticks roses fuchsia) de Blo t 21 (code PDB 2LM9 en saumon).

Il existe plusieurs méthodes de détermination des épitopes B des allergènes: par la résolution des structures des complexes anticorps-antigènes expérimentalement, par la méthode de phage display ou par mutagenèse dirigée. Cette dernière reste le moyen le plus efficace et le plus utilisé. En effet, plusieurs études ont montré que des substitutions par mutagenèse dirigée au niveau des résidus des épitopes B (généralement des acides aminés chargés présents à la surface de la protéine) peuvent avoir un impact très considérable sur sa IgE-réactivité en abolissant l'interaction avec les IgE spécifiques (IgEs) (Pomés, 2010). Ainsi, plusieurs épitopes des allergènes d'acariens ont été identifiés grâce à cette méthode. Citons ceux de Der p 2 (Mueller, 2001), de Blo t 5 (Chan *et al.*, 2008), de Blo t 21 (Tan *et al.*, 2012) et de Der f 13 (Chan *et al.*, 2006). Néanmoins, faire de la mutagenèse dirigée tout le long de la séquence des allergènes est une tâche lourde à mettre en place, que se soit en terme de coût ou en terme de temps. C'est pourquoi, une présélection des épitopes potentiels *in silico*, via l'utilisation des outils bioinformatiques de prédiction, est souhaitable et semble être une voie extrêmement prometteuse dans l'élaboration des vaccins hypoallergéniques.

La première partie de notre travail s'inscrit dans ce contexte. Dans ce présent chapitre, nous décrivons dans un premier temps l'analyse de séquences des allergènes des familles 5 et 21 dans le but d'étudier leur réactivité croisée en nous basant sur les épitopes de Blo t 5 et de Blo t 21 déterminés expérimentalement. Ensuite, nous présenterons la prédiction *in silico* de la position des épitopes B de l'allergène Der p 5. Ce travail contribue à l'étude de l'allergénicité des allergènes de ces deux familles (5 et 21) et aide à l'élaboration des variants hypoallergéniques et à la conception rationnelle des vaccins contre ces derniers.

*Chapitre II:*

*Méthodes*

## II.2 Méthodes

### II.2.1 Etude de la réactivité croisée entre les allergènes des familles 5 et 21

Afin d'étudier la réactivité croisée entre les allergènes de la famille 5 (Der p 5, Der f 5, Blo t 5, Sui m 5 et Lep d 5) et leurs homologues de la famille 21 (Der p 21, Der f 21 et Blo t 21), nous avons tout d'abord commencé par la mesure de leur similarité de séquence en réalisant des alignements par paires et multiples à l'aide des programmes bioinformatiques Blast2seq (Altschul et *al.*, 1997) et ClustalW2 (Larkin et *al.*, 2007), respectivement. Par ailleurs, par manque de données sur les allergènes Ale o 5, Tyr p 5 et Gly d 5, ces derniers n'ont pas été étudiés dans notre travail.

Les séquences des allergènes étudiés ont été collectées dans la base de données UniProtKB/SwissProt (<http://www.uniprot.org>). Leurs codes d'accès sont P14004 pour Der p 5, O96870 pour Blo t 5, A8B8I1 pour Der f 5, B2GM87 pour Sui m 5, Q9U5P2 pour Lep d 5, Q2L7C5 pour Der p 21, A8B8G7 pour Der f 21 et A7IZE9 pour Blo t 21.

#### II.2.1.1 Alignement de séquences

Dans un premier temps, des alignements locaux par paires ont été effectués entre ces séquences en utilisant le programme Blast2seq (Altschul et *al.*, 1997) afin d'estimer leurs pourcentages d'identité et de similarité de séquences en utilisant la matrice de score Blosum 62. Ces derniers sont définis, respectivement, comme étant le pourcentage des résidus identiques qui se trouvent dans les mêmes positions entre les deux séquences comparées et le pourcentage d'acides aminés pour lesquels le score d'alignement (selon notre matrice de score) est positif.

Nous avons ensuite utilisé le programme ClustalW2 avec ses paramètres par défauts (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>) (Larkin et *al.*, 2007) pour l'alignement multiple entre ces huit séquences. Cet alignement va nous permettre de discuter la conservation ou non des régions épitopes de Blo t 5 et de Blo t 21 dans les séquences des autres allergènes étudiés et donc essayer de comprendre la variabilité (forte ou faible) de la réactivité croisée entre ces allergènes des familles 5 et 21.

### II.2.2 Prédiction des épitopes de Der p 5

Afin de prédire les épitopes B sur la base de la structure 3D de l'allergène Der p 5, nous avons utilisé le serveur web du programme SEPPA (<http://lifecenter.sgst.cn/seppa/>, Sun et *al.*, 2009)

vu ses performances par rapport à d'autres programmes évalués par Xu et *al.* (2010). De plus, ce programme a pu prédire correctement la position des épitopes de Blo t 5 et de Blo t 21 déterminés expérimentalement (Chan et *al.*, 2008; Naik et *al.*, 2008; Tan et *al.*, 2012).

Ce programme se base sur le concept appelé "unité de patch triangulaire de résidus". Cette unité est introduite afin de mieux décrire le contexte spatial local en surface de l'antigène selon un critère de distance. Elle est définie sur l'ensemble de 3 résidus en surface dont les distances deux à deux sont au plus de 4 Å. Pour chaque résidu "r" en surface, SEPPA recherche ensuite toutes les unités de patches situées dans un rayon de 15 Å autour de ce résidu et lui attribue un score d'antigénicité. Ce dernier est fonction de la propension de chaque unité de patch à faire partie d'un épitope (selon la banque d'apprentissage de SEPPA) et d'un coefficient d'agrégation des patches autour du résidu considéré.

### II.2.2.1 Structures expérimentales des allergènes Der p 5, Blo t 5 et Blo t 21

Les structures 3D de Der p 5, Blo t 5 et de Blo t 21 ont été collectées de la *Protein Data Bank* (PDB, <http://www.pdb.org>), avec les codes PDB 3MQ1, 2JMH et 2LM9, respectivement.

La structure de Der p 5 a été résolue par cristallographie aux rayons-X. Cette structure a été trouvée sous forme d'hexamère, un triplet d'homo-dimères (chaînes A-B, chaînes C-D et chaînes E-F). L'évaluation de la qualité stéréochimique de ces dimères par PROCHECK (Laskowski et *al.*, 1993) montre que la géométrie des résidus (la distribution des angles dièdres phi ( $\phi$ ) et psi ( $\psi$ ) dans le graphique de Ramachandran) est favorable (95.4% des résidus se trouvent dans les régions favorables sur l'ensemble des chaînes). En revanche, l'analyse de la carte de densité électronique des dimères de Der p 5 montre que contrairement à ce qui a été observé dans chaînes A et B, les chaînes C, D, E et F contiennent plusieurs résidus qui présentent une résolution nettement moins bonne que celle de la moyenne des résidus du même type dans les structures déterminées à une résolution semblable. De plus, les deux chaîne A et B ont la plus basse valeur du facteur B (facteur d'agitation thermique), suivi par le dimère C-D puis le dimère E-F (Mueller et *al.*, 2010a). Ces observations montrent clairement que le dimère le plus représentatif de Der p 5 est celui composé des deux chaînes A et B (3MQ1\_AB). C'est pour cette raison que dans la suite de notre recherche, nous avons travaillé sur le monomère 3MQ1\_A et sur le dimère 3MQ1\_AB.

Les structures 3D de Blo t 5 et de Blo t 21 ont été déterminées par RMN. Le fichier PDB de Blo t 5 (2JMH, Naik et *al.*, 2008) contient 21 conformères qui correspondent à 20 solutions de structures possibles (mais très similaires) satisfaisant les contraintes de distance entre les atomes et une structure (le conformère n°21) qui est en fait la moyenne des ces 20 conformations. La première conformation (le conformère n°1) dans ce fichier est dite la structure représentative. Par ailleurs, dans le fichier PDB de Blo t 21 (2LM9, Tan et *al.*, 2012), il y a 20 conformères et celui ayant la plus basse énergie parmi eux est donné comme la conformation la plus représentative (le conformère n°1).

### II.2.3 Criblage des résidus épitopes majeurs de Der p 5 pour la validation expérimentale

En vue de déterminer les résidus majeurs qui pourraient jouer un rôle critique dans la fixation aux IgEs, nous avons proposé des mutations en "alanine" de quelques résidus prédits par SEPPA comme épitopes.

La sélection de ces résidus s'est basée sur leur nature et leur accessibilité au solvant. En effet, des études préalables ont montré que, par rapport aux résidus non épitopes, les résidus épitopes présentent une accessibilité au solvant plus élevée et sont surtout des résidus polaires et aromatiques (Rubinstein et *al.* 2008; Sun et *al.*, 2011).

L'accessibilité au solvant d'un résidu X dans la structure de nos allergènes est définie comme le rapport (en pourcentage) entre la surface accessible au solvant de ce résidu, calculée à l'aide du programme DSSP (Kabsch & Sander, 1983) et sa surface accessible maximale. Celle-ci est définie comme l'accessibilité au solvant du résidu X au sein d'un tripeptide Gly-X-Gly en conformation étendue (Rose et *al.*, 1985).

Le choix de la mutation de ces résidus en alanine, quant à lui, s'est porté sur (1) la base des données expérimentales de mutagenèse dirigée de Blo t 5 et de Blo t 21, et surtout (2) sur le fait que ce résidu soit préféré dans les structures sous forme d'hélice alpha comme celle de Der p 5 (code PDB 3MQ1; Mueller et *al.*, 2010a). Néanmoins, afin d'estimer l'impact de ces mutations sur la stabilité thermodynamique de la structure de Der p 5, nous avons calculé la différence d'énergie libre ( $\Delta\Delta G$ ) entre la protéine sauvage et la protéine mutée en utilisant le serveur web du programme PoPMuSiC2.1 (Dehouck et *al.*, 2011). Plusieurs travaux ont montré l'efficacité de ce programme dans la prédiction des changements de stabilité après mutation ponctuelle (Dehouck et *al.*, 2012).

Ce programme a été également testé sur les structures 3D de Blo t 5 (code pdb 2JMH) et celle de Blo t 21 (code pdb 2LM9) afin d'estimer les valeurs de " $\Delta\Delta G$ " après la mutation de leurs épitopes déterminés expérimentalement en alanine.

Afin d'avoir des résultats plus significatifs et représentatifs de la différence de l'orientation des chaînes latérales dans les différentes conformations de la structure de Blo t 5 et de Blo t 21, nous avons calculé la moyenne de l'accessibilité au solvant et la moyenne de " $\Delta\Delta G$ " données en chaque position dans les 20 conformères des structures 2JMH, 2LM9 de Blo t 5 et Blo t 21, respectivement.

#### **II.2.4 Visualisation des structures**

Le logiciel PyMOL (<http://pymol.sourceforge.net/>; DeLano, 2002) a été utilisé pour visualiser toutes les structures 3D des allergènes étudiés.

## *Chapitre II:*

### *Résultats et discussions*

### II.3 Résultats et discussions

Deux objectifs ont été tracés dans ce travail: (1) l'analyse de la similarité de séquence entre les allergènes d'acariens des familles 5 et 21 (Der p 5, Blo t 5, Der f 5, Sui m 5, Lep d 5, Der p 21, Der f 21 et Blo t 21) afin de discuter leur réactivité croisée sur la base de la connaissance des épitopes de Blo t 5 et de Blo t 21 déterminés expérimentalement; et (2) la prédiction des épitopes de l'allergène Der p 5 vu la disponibilité de sa structure 3D et son rôle déterminant dans les maladies respiratoires et l'eczéma. Ce travail s'effectue en collaboration avec le Centre d'Ingénierie des Protéines (CIP) de l'Université de Liège (ULg), pour permettre une validation expérimentale de nos prédictions.

#### II.3.1 Analyse de l'alignement de séquences

L'alignement de séquences par paires de nos allergènes a été réalisé avec le programme Blast2seq en utilisant la matrice de substitution Blosum62 (Henikoff et Henikoff, 1992). L'analyse des résultats montre que ces séquences sont très similaires (tableau n°8). Le pourcentage d'identité de séquence est supérieur à 40%, excepté pour Derp5/Lepd5, Der p 5/Der f 21, Der p 5/Der p 21, Der p 5/Blo t 21, Der f 5/Der f 21, Der f 5/Der p 21 et Der f 5/Blo t 21. Notons que les plus grands pourcentages d'identité et de similarité ont été trouvés entre Der p 5 et Der f 5 (77% et 90%, respectivement) et entre Der p 21 et Der f 21 (74% et 87%, respectivement).

**Tableau n°8:** Pourcentage d'identité (triangle supérieur de la matrice) et de similarité (triangle inférieur de la matrice) de séquences entre les allergènes étudiés. La longueur de l'alignement est donnée entre parenthèse.

	Blo t 5	Der p 5	Der f 5	Sui m 5	Lep d 5	Der f 21	Der p 21	Blo t 21
Blo t 5	/	44% (107)	47% (107)	53% (107)	48% (110)	44% (109)	44% (107)	43% (106)
Der p 5	71% (107)	/	<b>77%</b> (109)	51% (109)	37% (107)	33% (106)	30% (106)	30% (108)
Der f 5	74% (107)	90% (109)	/	53% (109)	42% (107)	37% (106)	35% (106)	35% (107)
Sui m 5	74% (107)	68% (109)	66% (109)	/	46% (107)	43% (105)	41% (108)	46% (105)
Lep d 5	64% (110)	59% (107)	61% (107)	64% (107)	/	50% (109)	50% (107)	47% (108)
Der f 21	67% (109)	57% (106)	60% (106)	66% (107)	71% (109)	/	<b>74%</b> (111)	40% (105)
Der p 21	64% (107)	55% (106)	59% (106)	63% (108)	69% (107)	87% (111)	/	40% (105)
Blo t 21	72% (106)	60% (108)	64% (107)	70% (105)	79% (108)	71% (105)	69% (105)	/

La réactivité croisée entre les allergènes est due à la conservation ou non des sites antigéniques (épitopes) qui se trouvent à la surface de ces allergènes. Rappelons que les anticorps de type IgE sont spécifiques à l'espèce (IgEs), néanmoins, si l'épitope d'un allergène

"A" est identique à celui d'un autre allergène "B", il peut y avoir une réactivité croisée entre ces deux allergènes (A et B) en fixant ces mêmes IgEs.

King & Spaugfort, 2000 ont suggéré que les identités de séquences supérieures à 70% étaient nécessaires pour entraîner une réactivité croisée significative entre les allergènes homologues. C'est le cas pour les allergènes du groupe 1 entre Der p 1, Der f 1 et Eur m 1 (Takai et *al.*, 2005), du groupe 2 entre Der p 2, Der f 2 et Eur m 2 (Smith et *al.*, 2001) et du groupe 10 entre Der p 10 et Lep d 10, qui présentent entre eux une forte réactivité croisée (Witteaman et *al.*, 1994). En nous basant sur les pourcentages d'identité calculés entre les séquences de ces allergènes (tableau n°8), nous pouvons suggérer que seuls les allergènes Der p 5 et Der f 5 et les allergènes Der p 21 et Der f 21 pourraient avoir une forte réactivité croisée entre eux. Cependant, aucune mesure expérimentale de la réactivité croisée entre ces allergènes n'a été testée jusqu'à présent (tableau 7).

L'alignement multiple de nos séquences réalisé à l'aide de ClustalW2 reflète le degré de conservation des résidus tout au long des séquences d'allergènes étudiés (figure n°13). Ici, nous nous sommes intéressés à comparer les séquences au niveau des épitopes déterminés expérimentalement pour Blo t 5 (Naik et *al.*, 2008; Chan et *al.*, 2008) et pour Blo t 21 (Tan et *al.*, 2012) avec leurs correspondants dans les séquences des autres allergènes (Der p 5, Der f 5, Sui m 5, Lep d 5, Der p 21 et Der f 21). La numérotation des résidus utilisée dans la suite de ce travail sera celle de notre alignement multiple de séquence (figure n°13).

Deux épitopes ont été identifiés dans Blo t 5 (Chan et *al.*, 2008; Naik et *al.*, 2008). Les résidus majeurs de ces épitopes (épitope-I E69, D74, E79 et E84 et épitope-II N39, K40, K47 et R50) sont représentés par des lettres rouges en gras dans l'alignement de séquences multiple (figure n°13). Par ailleurs, un seul épitope a été trouvé dans Blo t 21 (Tan et *al.*, 2012). Les résidus majeurs de cet épitope (épitope-I E69, D74, E79, E84, et D91) sont indiqués en lettres bleues en gras.

N° dans le PDB 3MQ1	34										40					45					50														
N° de l'alignement	1					5					10					15					20					25					30				
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	L	L	Y	L	Q					
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	L	F	Y	L	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	L	L	H	L	Q					
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	L	L	A	L	T					
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	L	A	H	L	S					
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	L	L	M	L	S					
Blo t 21	D	N	F	R	H	E	F	D	H	M	I	V	N	T	A	T	Q	R	F	H	E	I	E	K	F	L	L	H	I	T					
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	L	L	H	L	S					

N° dans le PDB 3MQ1	55					60					65					70					75					80				
N° de l'alignement	35					40					45					50					55					60				
Blo t 5	H	Q	L	D	E	L	N	E	N	K	S	K	E	L	Q	E	K	I	I	R	E	L	D	V	V	C	A	M	I	E
Der p 5	E	Q	I	N	H	F	E	E	K	P	T	K	E	M	K	D	K	I	V	A	E	M	D	T	I	I	A	M	I	D
Der f 5	H	Q	I	N	T	F	E	E	N	P	T	K	E	M	K	E	Q	I	L	G	E	M	D	T	I	I	A	L	I	D
Sui m 5	E	Q	I	A	H	L	E	Q	S	K	N	K	E	E	K	E	K	I	V	R	E	L	E	T	I	I	A	L	I	S
Lep d 5	H	E	V	T	E	L	E	K	S	K	S	K	E	L	K	A	Q	I	L	R	E	I	S	I	G	L	L	M	L	S
Der f 21	E	Q	Y	K	E	L	E	K	T	K	S	K	E	L	K	E	Q	I	L	R	E	L	T	I	A	E	N	Y	L	R
Blo t 21	H	E	V	D	D	L	E	K	T	G	N	K	D	E	K	A	R	L	L	R	E	L	T	V	S	E	A	F	I	E
Der p 21	E	E	Y	K	E	L	E	K	T	K	S	K	E	L	K	E	Q	I	L	R	E	L	T	I	G	E	N	F	M	K

N° dans le PDB 3MQ1	85					90					95					100					105					110				
N° de l'alignement	65					70					75					80					85					90				
Blo t 5	G	A	Q	G	A	L	E	R	E	L	K	R	T	D	L	N	I	L	E	R	F	N	Y	E	E	A	Q	T	L	S
Der p 5	G	V	R	G	V	L	D	R	L	M	Q	R	K	D	L	D	I	F	E	Q	Y	N	L	E	M	A	K	K	S	G
Der f 5	G	V	R	G	V	L	N	R	L	M	K	R	T	D	L	D	I	F	E	R	Y	N	V	E	I	A	L	K	S	N
Sui m 5	G	S	H	D	V	L	E	R	E	L	K	R	T	D	L	D	I	L	E	R	Y	N	F	E	S	A	L	K	I	G
Lep d 5	S	A	K	G	H	F	E	R	E	L	K	R	A	D	L	N	L	A	E	K	F	N	F	E	S	A	L	S	T	G
Der f 21	G	A	L	K	F	M	Q	Q	E	A	K	R	T	D	L	N	M	F	E	R	Y	N	F	E	T	A	V	S	T	I
Blo t 21	G	S	R	G	Y	F	Q	R	E	L	K	R	T	D	L	D	L	L	E	K	F	N	F	E	A	A	L	A	T	G
Der p 21	G	A	L	K	F	F	E	M	E	A	K	R	T	D	L	N	M	F	E	R	Y	N	Y	E	F	A	L	E	S	I

N° dans le PDB 3MQ1	115					120					125					105					110								
N° de l'alignement	95					100					105					110													
Blo t 5	K	I	L	L	K	D	L	K	E	T	E	Q	K	V	K	D	I	Q	T	Q	-	-	-	-	-	-	-	-	-
Der p 5	D	I	L	E	R	D	L	K	K	E	E	A	R	V	K	K	I	E	V	-	-	-	-	-	-	-	-	-	-
Der f 5	E	I	L	E	R	D	L	K	K	E	E	Q	R	V	K	K	I	E	V	-	-	-	-	-	-	-	-	-	-
Sui m 5	A	I	L	V	R	D	L	K	A	A	E	A	K	V	K	A	I	N	V	H	A	-	-	-	-	-	-	-	-
Lep d 5	A	V	L	H	K	D	L	T	A	A	L	A	T	K	V	K	A	I	E	T	K	-	-	-	-	-	-	-	-
Der f 21	E	I	L	V	K	D	L	A	E	L	A	K	K	V	K	A	V	K	S	D	D	-	-	-	-	-	-	-	-
Blo t 21	D	L	L	L	K	D	L	K	A	L	Q	K	R	V	Q	D	S	E	-	-	-	-	-	-	-	-	-	-	-
Der p 21	K	L	L	L	K	D	L	D	E	L	A	K	K	V	K	A	V	N	P	D	E	Y	Y	-	-	-	-	-	-

X Résidus majeurs dans les épitopes de Blo t 5 (Chan et al., 2008; Naik et al., 2008)      X Résidus majeurs dans l'épitope de Blo t 21 (Tan et al., 2012)  
X Epitopes de Blo t 5 prédits par SEPPA      X Epitopes de Blo t 21 prédits par SEPPA  
X Epitopes de Der p 5 prédits par SEPPA

**Figure n°13:** Alignement multiple de séquences des allergènes des familles 5 et 21 étudiés. Positions des épitopes déterminés expérimentalement de Blo t 5 et de Blo t 21 et ceux prédits par le programme SEPPA de Blo t 5, de Blo t 21 et de Der p 5. La première ligne contient la numérotation des résidus dans la structure cristallographique de Der p 5 (3MQ1). La ligne suivante correspond à la numérotation de l'alignement de séquences. Blo t : *Blomia tropicalis*, Der p : *Dermatophagoides pteronyssinus*, Der f : *Dermatophagoides farinae*, Sui m : *Suidasia medanensis*, Lep d : *Lepidoglyphus destructor*.

D'après la littérature, une faible réactivité croisée a été trouvée entre les allergènes Der p 5/Blo t 5, Der p 5/Lep d 5, Der p 5/Der p 21, Der p 5/Der f 21, Der p 5/Blo t 21, Blo t 5/Der f 5, Blo t 5/Blo t 21, Blo t 5/Der p 21, Blo t 5/Der f 21, Blo t 21/Der p 21, Blo t 21/Der f 21 et Der p 21/Lep d 5. En outre, une forte réactivité croisée a été observée entre Sui m 5 et Blo t 5 et entre Sui m 5 et Der f 5.

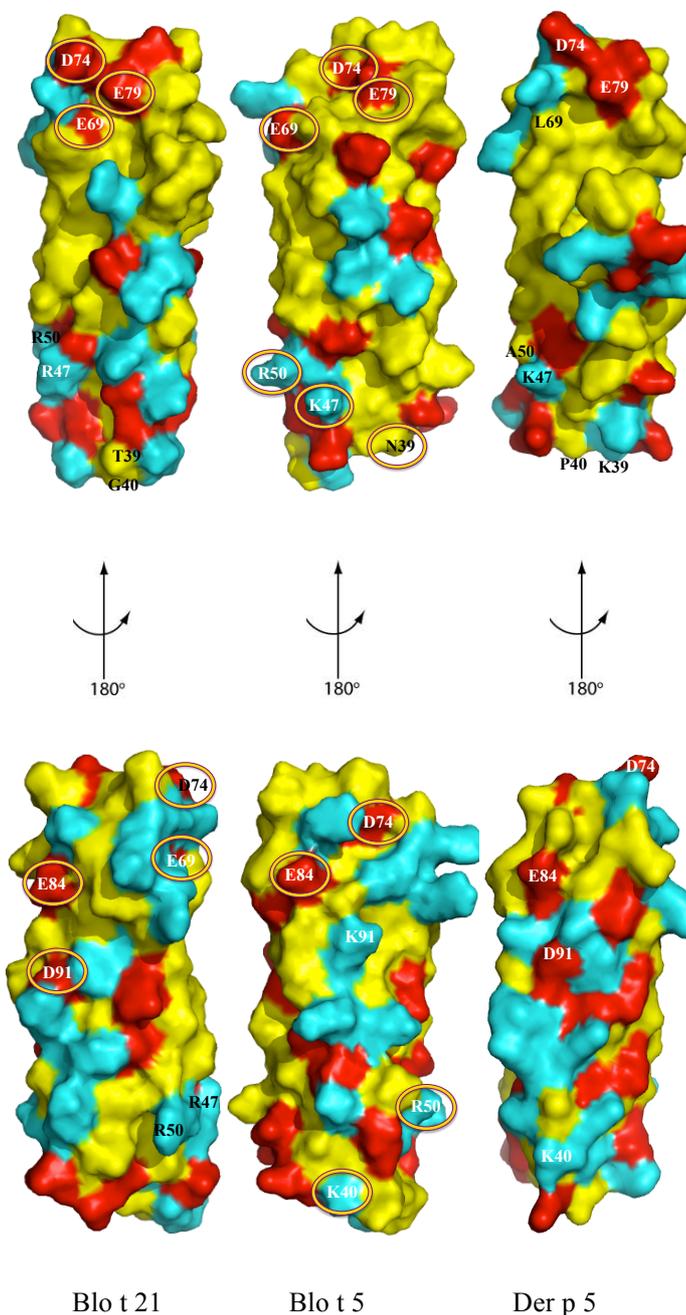
L'épitope I (composé des résidus majeurs E69, D74, E79 et E84) est le plus important dans Blo t 5 (Chan et al., 2008). Ces quatre résidus majeurs sont conservés dans toutes les

séquences des allergènes des familles 5 et 21 étudiés (figure n°13), excepté pour le résidu glutamate en position 69 (E69) qui est substitué par un résidu hydrophobe, leucine (L), dans Der p 5 et Der f 5. Ce qui est intéressant, c'est que l'épitope-I de Blo t 21 identifié par mutagenèse dirigée par Tan et *al.* (2012) est constitué de ces quatre résidus majeurs de Blo t 5, plus un résidu aspartate dans la position 91 (D91). Ce résidu est conservé dans Der p 5, mais substitué par un résidu de même charge, le glutamate (E91), dans Der f 5 et Der f 21 et par un résidu de charge opposée, la lysine (K91), dans Blo t 5 et Der p 21. Par contre dans Sui m 5 et Lep d 5, ce résidu (E91) est remplacé par une alanine (A91).

L'épitope-II de Blo t 5 déterminé par Naik et *al.* (2008) est composé des quatre résidus majeurs N39, K40, K47 et R50. Ces résidus ne sont pas aussi bien conservés que l'épitope-I dans toutes les séquences des allergènes étudiés (figure n°13). En effet, le résidu N39 est conservé uniquement dans Der f 5 et remplacé par une lysine dans Der p 5 (K39), par une sérine (S39) dans Sui m 5 et Lep d 5 et par une thréonine (T39) dans les allergènes de la famille 21 (Der p 21, Der f 21 et Blo t 21). Le résidu K40 est conservé dans Sui m 5, Lep d 5, Der f 21 et Der p 21 et substitué par une proline (P40) dans Der p 5 et Der f 5 et par une glycine (G40) dans Blo t 21. Par ailleurs, le résidu K47 est conservé dans Der p 5 et dans Sui m 5 uniquement, mais substitué par un résidu de même charge positive, l'arginine (R47), dans Blo t 21 et par une glutamine (Q47) dans Der f 5, Lep d 5, Der f 21 et Der p 21. Le résidu R50, quant à lui, est conservé dans la majorité de ces allergènes (Sui m 5, Lep d 5, Der f 21, Blo t 21 et Der p 21) mais remplacé par une alanine (A50) dans Der p 5 et par une glycine (G50) dans Der f 5. Ces substitutions par des résidus présentant des propriétés physico-chimiques différentes au niveau des résidus épitopes majeurs peuvent avoir un impact sur l'interaction de ces derniers avec les IgE, ce qui pourrait expliquer en partie leur faible réactivité croisée.

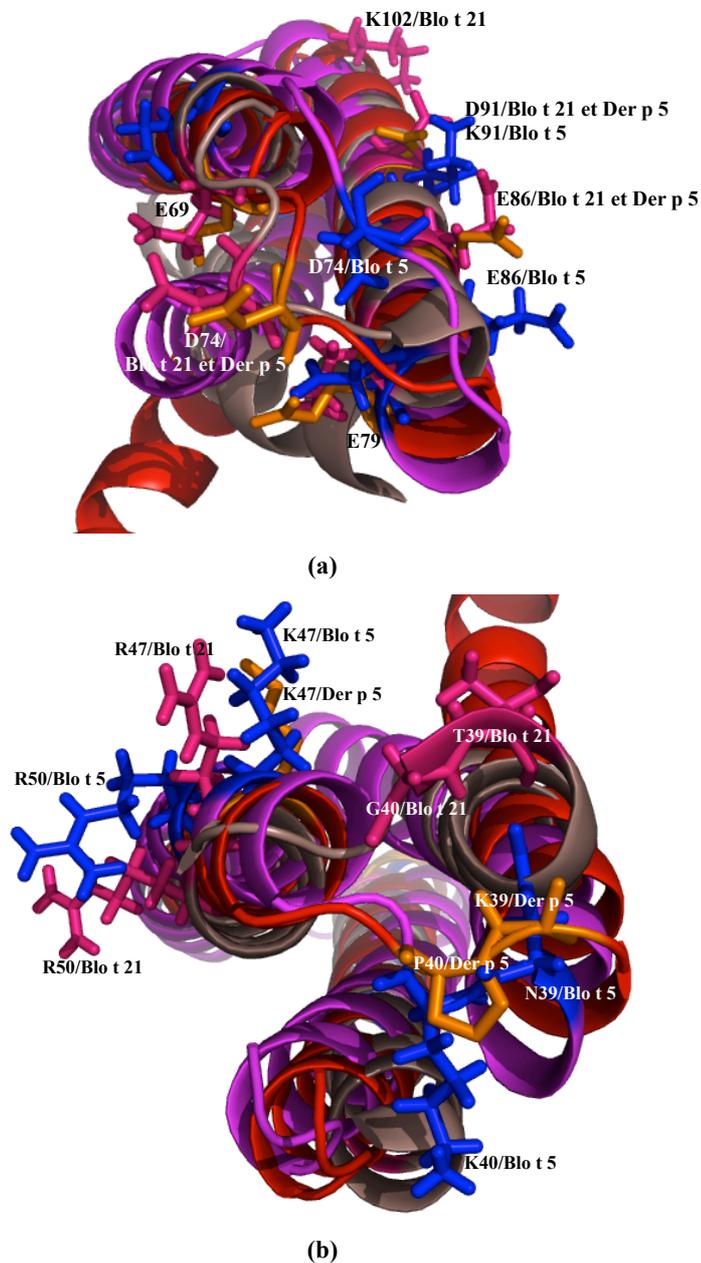
Seul l'allergène Sui m 5 présente une conservation de trois résidus sur les quatre de cet épitope de Blo t 5 (K40, K47 et R50) et une conservation de la totalité des quatre résidus de l'épitope I (E69, D74, E79 et E84). Ceci pourrait expliquer la forte réactivité croisée entre ces deux allergènes décrite par Reginald et *al.* (2004). Cependant, une forte réactivité est également observée entre Sui m 5 et Der f 5, mais en se basant sur l'alignement de séquence avec les épitopes connus de Blo t 5 et de Blo t 21, nous remarquons que seul l'épitope-I est conservé entre ces allergènes. Cela nous laisse penser que la réactivité croisée est probablement due à cet épitope et/ou à l'existence d'autres épitopes non encore identifiés dans

ces allergènes. Par ailleurs, une faible réactivité croisée est observée entre Blo t 5/Blo t 21, Blo t 5/Der p 5 et Blo t 21/Der p 5 (tableau n°7). En nous focalisant sur les structures de Blo t 21, Blo t 5 et de Der p 5 (2LM9, 2JMH et 3MQ1\_A, respectivement) nous remarquons que la distribution des résidus chargés est différente entre ces trois allergènes (figure n° 14).



**Figure n° 14:** Distribution des résidus chargés (K, R et H en cyan et E et D en rouge) dans les structures (en jaune) de Blo t 21 (code PDB 2LM9), de Blo t 5 (code PDB 2JMH) et de Der p 5 (code PDB 3MQ1\_A). Les positions des résidus épitopes majeurs de Blo t 5 et de Blo t 21 déterminés expérimentalement sont entourées par des cercles oranges.

Pour comparer les résidus des épitopes de Blo t 5, Blo t 21 avec leurs correspondant dans Der p 5, nous avons fait une analyse de leur alignement de séquence (figure n°13) et de leurs structures 3D (figures n°14 et 15):



**Figure n°15:** Superposition de la structure au niveau des épitopes I (a) et II (b) de Blo t 5 (en sticks bleus), de Blo t 21 (en sticks roses) et leurs correspondants dans Der p 5 (en sticks oranges). Les structures 3D de Blo t 5 (code PDB 2JMH), de Blo t 21 (code PDB 2LM9) et de Der p 5 (code PDB 3MQ1) sont représentées en cartoon magenta, saumon et rouge, respectivement.

## ↻ entre Blo t 5 et Blo t 21

- Au niveau de l'épitope-I les quatre résidus majeurs de Blo t 5 sont conservés sur les cinq résidus identifiés dans Blo t 21. Le résidu majeur D93 de Blo t 21 est substitué par une charge opposée (K93) dans Blo t 5. La position de ces résidus majeurs (E69, D74, E79, E84) dans leurs structures est similaire (tenant compte de la flexibilité de la boucle et des chaînes latérales dans les 20 conformères de ces structures), sauf pour le résidu E69 qui est un peu enfoui dans la structure de Blo t 21 (2LM9).

- Au niveau de l'épitope-II de Blo t 5, un seul résidu (R50) sur quatre est conservé dans Blo t 21. Par ailleurs, une charge est également conservée par la substitution de K47 de Blo t 5 par une arginine (R47) dans Blo t 21. Les autres résidus N39 et K40 de Blo t 5 sont remplacés par une thréonine (T39) et une glycine (G40), respectivement. La position de ces résidus dans la structure de ces allergènes est variable (figures 13 et 14).

## ↻ entre Blo t 5 et Der p 5:

- Au niveau de l'épitope-I, il y a trois résidus sur quatre qui sont conservés. Le résidu E69 est substitué par une leucine dans Der p 5 (L69). La position de ces trois résidus est similaire dans les structures de Blo t 5 et de Der p 5 (2JMH et 3MQ1\_A, respectivement).

- Au niveau de l'épitope-II, trois résidus sur quatre sont conservés (E69 de Blo t 5 est substitué par une leucine dans Der p 5). La position de ces résidus dans leurs structures est similaire (résidus exposés en surface).

## ↻ entre Blo t 21 et Der p 5

- Au niveau de l'épitope-I, quatre résidus (D74, E79, E84 et D91) sur cinq sont conservés. Le résidu épitope majeur E69 de Blo t 21 est substitué par une leucine dans Der p 5 (L69). La position de ces quatre résidus ainsi que l'orientation de leurs chaînes latérales sont très similaires dans les structures de Der p 5 et de Blo t 21 (3MQ1\_A et 2LM9, respectivement).

Toutes ces observations nous amènent à trois hypothèses pour expliquer la faible réactivité croisée entre ces allergènes. (1) La faible réactivité croisée observée entre Blo t 5, Blo t 21 et Der p 5 est probablement due à l'épitope I, (2) la substitution des autres résidus qui se trouvent dans cet épitope (les résidus adjacents aux résidus épitopes majeurs) pourraient également jouer un rôle dans la spécificité de la reconnaissance des IgE, et/ou (3) la possibilité d'existence d'autres sites antigéniques au sein de ces allergènes.

II.3.2 Prédiction des épitopes de Der p 5

L'analyse de la structure 3D et de l'alignement de séquences et de structures indiquent que les résidus de l'épitope-I de Blo t 5 et de Blo t 21 sont fortement conservés dans Der p 5. De plus, du point de vue son allergénicité, il est intéressant de chercher les épitopes de Der p 5 qui pourraient servir à la conception des variants hypoallergéniques de cet allergène.

Il existe plusieurs méthodes de prédiction des épitopes B, celles qui se basent sur la séquence des allergènes ou celles qui se basent sur leurs structures. Ces dernières méthodes semblent être les plus performantes vu que 90% des épitopes sont conformationnels (liés à leurs structures 3D). Xu *et al.* (2010) ont évalué la performance de quelques programmes qui se basent sur la structure 3D des allergènes pour la prédiction des épitopes B: "SEPPA" (Sun *et al.*, 2009), "CEP" (Kulkarni-Kale *et al.*, 2005), "DiscoTope" (Andersen *et al.*, 2006), "ElliPro" (Ponomarenko *et al.*, 2008), "PEPOP" (Moreau *et al.*, 2008) et "BEpro/PEPITO" (Sweredoski & Baldi, 2008). Ces auteurs ont montré que SEPPA (Sun *et al.*, 2009) est l'outil de prédiction le plus performant. De plus, nous avons testé ce programme sur la structure de Blo t 5 (code PDB 2JMH) et celle de Blo t 21 (code PDB 2LM9) pour lesquels les épitopes ont été déterminés expérimentalement (Chan *et al.*, 2008; Naik *et al.*, 2008; Tan *et al.*, 2012). La majorité des résidus épitopes majeurs de Blo t 5 et de Blo t 21 ont été prédit correctement par SEPPA (4 résidus sur 4 dans l'épitope-I, et 2 résidus sur 4 dans l'épitope-II de Blo t 5 et 4 résidus sur 5 dans l'épitope I de Blo t 21) (figure n°15). Les résultats de la prédiction des épitopes de Der p 5 avec ce programme sont résumés dans la figure n° 16.

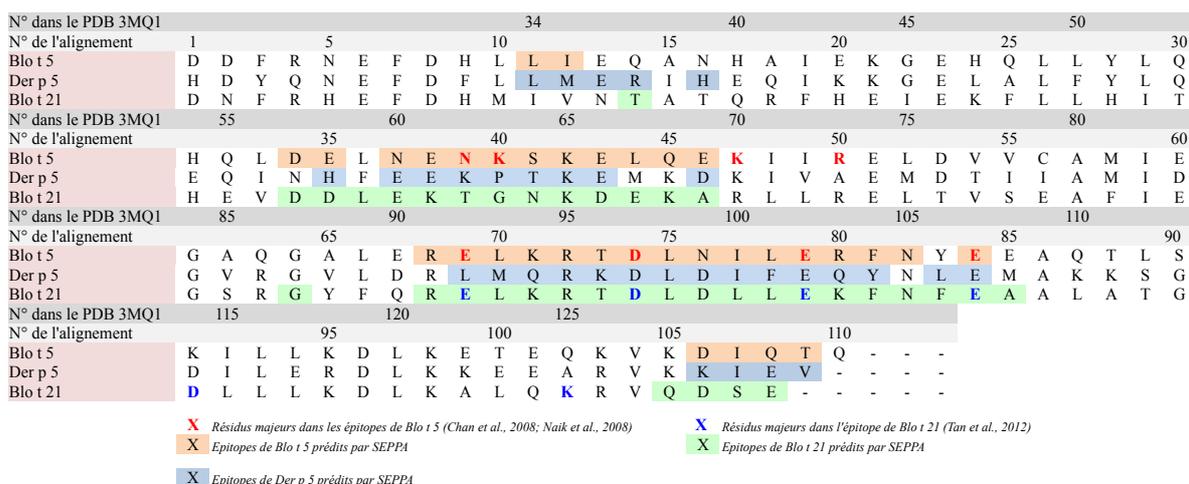


Figure n°16: Alignement de séquence multiple entre Blo t 5, Der p 5 et Blo t 21 et positions de leurs épitopes B prédits par le programme SEPPA (cet alignement est repris de la figure n°13).

Quatre sites antigéniques ont été prédits dans Der p 5, Blo t 5 et Blo t 21 par le programme SEPPA:

1. Epitope-Nt : quelques résidus de la première hélice ont été prédits comme épitopes. Ceci correspond aux résidus L11 et I12 pour Blo t 5, T14 pour Blo t 21 et L11, M12, E13, R14 et H16 pour Der p 5.
2. Epitope-II (selon notre convention de nomination des épitopes d'après leur importance): de la position 35 à 46 de notre alignement de séquences et qui correspond partiellement à l'épitope de Blo t 5 identifié par Naik *et al.* (2008);
3. Epitope-I (selon notre convention de nomination des épitopes d'après leur importance): de la position 62 à 85 de notre alignement de séquences et qui correspond à l'épitope identifié par Chan *et al.* (2008) pour Blo t 5 et par Tan *et al.* (2012) pour Blo t 21.
4. Epitope-Ct: quelques résidus de la l'hélice  $C_{terminale}$  ont également été prédits comme épitopes.

Ce sont essentiellement les résidus de l'épitope-I qui sont conservés entre Blo t 5, Blo t 21 et Der p 5. Aucune information n'a été trouvée sur l'importance de la première hélice de ces allergènes dans leur allergénicité, contrairement à la position des quelques résidus dans l'hélice  $C_{terminale}$ . En effet, Tan *et al.* (2012) ont montré que la mutation de K102 réduit significativement la fixation des IgEs de Blo t 21. Ce résidu étant loin de l'épitope conformationnel de cet allergène, il pourrait faire partie d'un autre épitope de Blo t 21. De plus, selon Flicker *et al.* (2006), la position des épitopes dans la partie  $C_{terminale}$  des allergènes devrait faciliter la réticulation du récepteur FcεRI via le complexe IgE-antigène et joue un rôle primordial dans leur allergénicité en stimulant la dégranulation des mastocytes ou des basophiles et la libération de médiateurs de l'inflammation (Siraganian, 2003).

### III.3.2.1 Criblage des résidus épitopes majeurs de Der p 5 pour la validation expérimentale

Les résidus exposés en surface sont un critère important pour la détermination des sites antigéniques des allergènes (Bui *et al.*, 2007). En effet, cette exposition leur permet d'interagir avec le paratope des anticorps pour former le complexe antigène-anticorps. De plus, Sun *et al.* (2011) ont étudié les différences entre les résidus épitopes et les non épitopes en surface de l'antigène. Ils ont trouvé que les résidus épitopes présentent une accessibilité au solvant plus élevée, et sont surtout des résidus polaires et aromatiques. Plusieurs travaux ont montré

également l'importance des résidus chargés en surface dans la fixation des IgEs. Des mutations dirigées sur ces résidus chargés ont permis d'identifier les épitopes de différents allergènes d'acariens: Der p 2 (Mueller & *al.*, 2001), Der f 13 (Chan et *al.*, 2006), Blo t 5 (Chan et *al.*, 2008) et Blo t 21 (Tan et *al.*, 2012). Tous ces allergènes présentent des pourcentages élevés de résidus chargés en surface. L'abondance de ces résidus chargés tout au long des structures des allergènes peut expliquer que la grande solubilité de ces allergènes pourrait favoriser leurs propriétés allergéniques.

La sélection des résidus majeurs qui pourraient jouer un rôle crucial dans la fixation des IgEs de Der p 5 s'est donc portée sur deux critères: (1) la nature du résidu, en choisissant les résidus chargés (K, R, H, D, E) et les résidus polaires (N, Q, S et T) et le résidu polaire aromatique tyrosine (Y) et (2) l'accessibilité au solvant des résidus épitopes prédits par SEPPA.

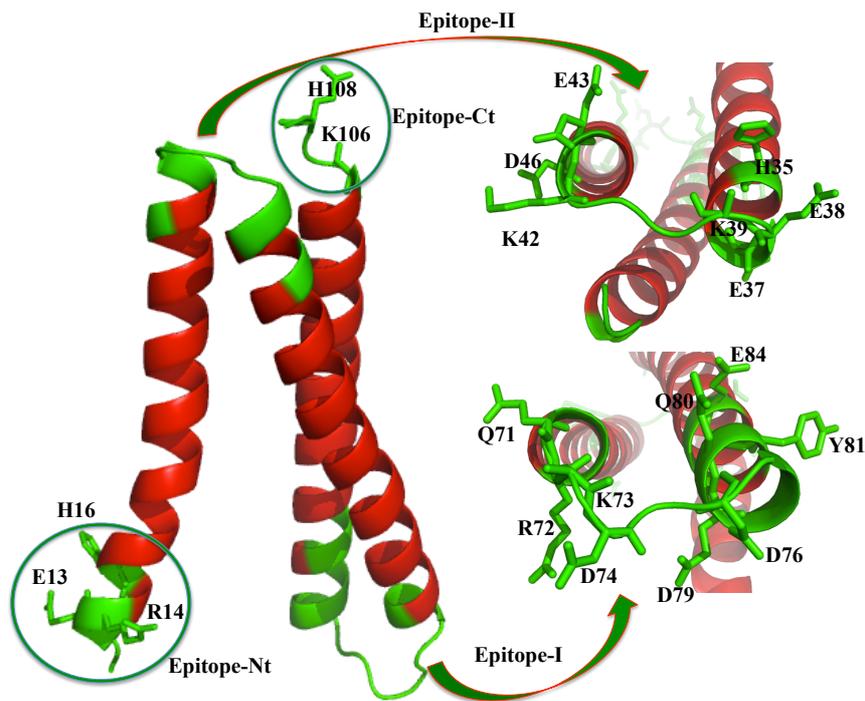
Afin de choisir une valeur seuil de l'accessibilité au solvant des résidus des épitopes, nous avons calculé avec DSSP la surface accessible au solvant des résidus majeurs des épitopes de Blo t 5 et de Blo t 21 déterminés expérimentalement. L'accessibilité au solvant donnée dans nos résultats correspond à la moyenne calculée pour ces résidus majeurs de des épitopes dans l'ensemble des 20 conformères de leurs structures RMN +/- écart-type ( $\langle \%acc \rangle \pm \sigma$ ) (tableau n°9). Les résultats trouvés montrent que l'accessibilité moyenne au solvant est supérieure à 25%, excepté pour le résidu E69 de Blo t 21 qui présente une accessibilité moyenne de 15%. Cependant, c'est probablement sa nature et sa position dans la structure, par rapport aux autres résidus majeurs, qui ont fait que cet épitope soit important pour Blo t 21.

En outre, l'accessibilité au solvant des résidus chargés, polaires et aromatiques prédits par SEPPA dans l'ensemble des deux chaînes A et B de Der p 5 (3MQ1\_AB) est supérieure à 25% (tableau n°9). Ceci montre la forte exposition de ces résidus dans la structure de Der p 5.

**Tableau n° 9:** Valeurs de la moyenne de l'accessibilité au solvant +/- écart-type (<%acc.> +/-  $\sigma$ ) des résidus épitopes sauvages de Blo t 5 et de Blo t 21 déterminés expérimentalement, et de l'accessibilité au solvant (acc.) des résidus épitopes des deux chaînes A et B de Der p 5 prédits par SEPPA. La valeur moyenne de la différence d'énergie libre (< $\Delta\Delta G$ >) +/- écart-type ((< $\Delta\Delta G$ >) +/-  $\sigma$ ) pour Blo t 5 et Blo t 21 et la valeur de la différence d'énergie libre pour Der p 5 ( $\Delta\Delta G$ ) après leur mutation en alanine sont également fournies

Epitopes	Résidu muté	<%acc.> +/- $\sigma$ (%)		< $\Delta\Delta G$ > +/- $\sigma$ (kcal/mol)	
<b>Blo t 5</b>					
Epitope II	N39A	59,27±4,88		/	
	K40A	36,09±2,51		/	
	K47A	59,85±4,15		/	
	R50A	65,31±5,93		/	
Epitope I	E69A	56,09±9,43		0,06±0,15	
	D74A	54,6±30,4		0,03±0,31	
	E79A	30,55±5,21		0,70±0,28	
	E84A	57,03±10,2		0,26±0,31	
<b>Blo t 21</b>					
Epitope I	E69A	15,18±4,87		0,66±0,21	
	D74A	65,61±6,27		0,06±0,15	
	E79A	34,65±10,93		0,4±0,47	
	E84A	50,26±7,3		0,22±0,19	
	D91A	26,9±7,29		0,27±0,25	
Epitope-Ct	K102	54,99±6,17		0,14±0,2	
<b>Der p 5</b>					
Epitope-Nt		Chaîne A acc. (%)	Chaîne B acc. (%)	Chaîne A $\Delta\Delta G$ (kcal/mol)	Chaîne B $\Delta\Delta G$ (kcal/mol)
	L11A	80,79	24,6	0,12	0,05
	M12A	66,37	20,24	0	0,03
	E13A	81,63	46,9	-0,28	-0,3
	R14A	73,44	19,4	-0,1	-0,13
	H16A	45,43	44,44	0,34	-0,02
Epitope II	H35A	57,78	44,36	-0,33	-0,32
	E37A	42,96	48,87	0,24	0,2
	E38A	86,47	54,51	0,04	-0,1
	K39A	25,59	26,42	1,3	1,24
	P40A	33,38	36,22	1,26	1,13
	T41A	38,69	35,41	1,24	1,26
	K42A	53,59	37,27	0,22	0,84
	E43A	77,34	68,3	0,33	0,14
	D46A	45,37	47,47	-0,17	-0,41
Epitope I	L69A	31,07	8,63	0,91	1
	M70A	29,5	32,37	1,07	0,67
	Q71A	81,78	69,88	-0,17	-0,06
	R72A	63,67	34,11	-0,19	-0,55
	K73A	41,63	54,84	0,73	-0,09
	D74A	97,04	46,63	0,36	0,39
	L75A	19,68	12,34	2,08	1,86
	D76A	52,3	28,99	0,79	0,96
	I77A	45,86	45,76	0,33	0
	F78A	23,79	31,87	1,86	0,13
	E79A	53,17	10,47	0,13	0,25
	Q80A	50,72	40,89	0,03	0,18
	Y81A	53,63	27,87	0,38	0,18
	L83A	11,91	12,77	1,74	1,52
Epitope-Ct	E84A	51,56	51,02	-0,05	0,04
	K106A	34,54	42,07	0,92	0,32
	I107A	22,65	26,24	1,27	2,05
	E108A	95,6	69,82	-0,14	-0,06
	V109A	69,3	255,12	0,18	0,58

Nous avons donc retenu les résidus E13, R14, H16 dans l'épitope-N<sub>terminal</sub>, les résidus H35, E37, E38, K39, K42, E43 et D46 dans l'épitope-II, les résidus Q71, R72, K73, D74, D76, E79, Q80, Y81 et E84 dans l'épitope-I et les résidus K106 et E108 dans l'épitope-C<sub>terminal</sub> pour la validation expérimentale (figure n°17).



**Figure n° 17:** Position des épitopes prédits de Der p 5 par le programme SEPPA et des résidus majeurs proposés, d'après nos critères de sélection, pour la validation expérimentale.

Chan *et al.* (2008) et Tan *et al.* (2012) ont déterminé les épitopes de Blo t 5 et de Blo t 21, respectivement, par la méthode de mutagenèse dirigée en faisant des mutations en alanine sur tous les glutamates (E) et aspartates (D) tout au long de la séquence de ces deux allergènes. Certains de ces résidus ont été identifiés comme résidus majeurs des épitopes selon la forte diminution de la fixation aux IgE observée après leur mutation. Cependant, d'autres mutations n'ont pas pu être exprimées expérimentalement, alors que d'autres ont été éliminées pour le test de fixation aux anticorps suite aux changements conformationnels que ces mutations induisent au sein de ces allergènes (perte d'une partie de la structure secondaire, par exemple). Ces changements conformationnels sont probablement dus à la déstabilisation de la structure de ces allergènes suite à la mutation en alanine en certaines positions de la structure. En effet, l'effet de la mutation en alanine peut être double: soit elle déstabilise fortement la structure de

l'allergène au point que la structure s'en voit altérée. Dans ce cas, il est impossible de déterminer si le résidu muté est situé dans un épitope car la perte d'interaction avec l'anticorps peut simplement être due à une modification de la structure de l'allergène. Soit la mutation ne modifie pas la structure de l'allergène. Dans ce cas, une perte d'interaction avec l'anticorps suggère que le résidu muté est situé dans un épitope. C'est ce dernier cas qui est recherché.

C'est pour cette raison que nous avons estimé, en utilisant le programme PoPMuSiC, le changement de stabilité thermodynamique suite à une mutation en alanines des résidus épitopes prédits par SEPPA pour Der p 5 afin de ne sélectionner que les positions dont la mutation en alanine ne déstabilise pas de manière importante la structure native de cet allergène. PoPMuSiC est basé sur l'hypothèse qu'une mutation ponctuelle ne modifie pas significativement la conformation de la chaîne principale de la protéine. Après chaque mutation, PoPMuSiC calcule la valeur de la différence d'énergie libre ( $\Delta\Delta G$ ) entre la protéine sauvage et la protéine mutée. Plus la valeur de  $\Delta\Delta G$  est négative, plus la mutation est considérée comme stabilisante dans cette position. Cependant, nous pouvons avoir une mutation déstabilisante (avec des valeurs de  $\Delta\Delta G$  positives) dans une certaine mesure qui ne modifie pas la structure de l'allergène. La question qui se pose dans ce cas est jusque dans quelle mesure peut-on déstabiliser l'allergène sans qu'il y ait perte de structure (ou dit autrement, jusqu'à quelle valeur du  $\Delta\Delta G$  positive peut-aller). Afin d'apporter une piste de réponse à cela, nous avons tout d'abord testé PoPMuSiC sur l'épitope-I de Blo t 5 et de Blo t 21 déterminés expérimentalement par mutagenèse dirigée en alanine afin d'estimer l'étendue des valeurs de  $\Delta\Delta G$  qui à priori ne déstabilisent pas la structure de ces allergènes après cette mutation. La valeur de  $\Delta\Delta G$  donnée dans nos résultats est la moyenne des ( $\langle \Delta\Delta G \rangle$ ) +/- écart-type calculée aux positions des résidus majeurs des épitopes dans les 20 conformères pour Blo t 5 et Blo t 21. La plus grande valeur de  $\Delta\Delta G$  moyenne ( $\langle \Delta\Delta G \rangle$ ) +/- écart-type dans l'épitope-I de Blo t 5 est de  $0,70 \pm 0,28$  kcal/mol. Cette valeur est proche de celle calculée après mutations des résidus majeurs de l'épitope-I dans Blo t 21 ( $0,66 \pm 0,21$  kcal/mol).

Nous avons ensuite calculé la valeur de la différence d'énergie libre ( $\Delta\Delta G$ ) dans les deux chaînes A et B de Der p 5 (3MQ1\_AB) au niveau des résidus épitopes prédits par SEPPA et sélectionnés selon nos critères (représentés avec des cases rouges dans le tableau n°9). Le résultat montre que les  $\Delta\Delta G$  calculés sur ces deux chaînes sont inférieurs à 1 kcal/mol après mutation de tous ces résidus en alanine, sauf pour la lysine en position 39 (K39) ayant une valeur de 1,3 kcal/mol dans la chaîne A et de 1,24 kcal/mol dans la chaîne B. Ces résultats

suggèrent que la mutation en alanine de ces résidus n'est pas trop défavorable dans ces positions en terme de la stabilité thermodynamique de la structure de Der p 5.

### **II.3.2.2 Comparaison des épitopes prédits de Der p 5 avec leurs correspondants dans la séquence de Der f 5**

L'alignement de séquence entre Der p 5 et Der f 5 montre que la région de l'épitope-I est fortement conservée: 12 résidus sur 15 sont identiques (L69, M70, R72, D74, L75, D76, I77, F78, E79, Y81, et E84) et les trois résidus Q71, K73 et L83 sont substitués par les résidus K, T et V, respectivement dans Der f 5. Seulement six résidus sur douze sont conservés entre Blo t 5/Blo t 21 et Der p 5 dans cet épitope-I. Ces observations pourraient renforcer l'hypothèse de l'existence d'une forte réactivité croisée entre les deux allergènes Der p 5 et Der f 5. La résolution de la structure de Der f 5 apportera plus de lumière quant à la réactivité croisée de ces allergènes des familles 5 et 21.

*Chapitre II:*

*Conclusions*

## II.4 Conclusions

La création de variants hypoallergéniques ayant une immunogénicité conservée mais un potentiel allergénique réduit, est un véritable enjeu de santé. En effet, ces variants sont toujours capables de stimuler le système immunitaire (cellules lymphocytaires T) mais sont dépourvus d'une réactivité IgE et donc incapables d'activer la réponse humorale IgE-dépendante et d'induire les signes cliniques de l'allergie. Ce mécanisme vise à l'acquisition de la tolérance immunologique, par la production et la stimulation de cellules T régulatrices. Ces variants sont donc fortement recherchés pour le traitement préventif ou thérapeutique de maladies allergiques. Les objectifs de ce chapitre étaient tracés dans ce contexte.

Nous nous sommes tout d'abord intéressés à l'étude de l'allergénicité des allergènes des familles 5 et 21 en étudiant leur réactivité croisée. Par la suite, nous avons prédit, *in silico*, les épitopes de Der p 5, un des allergènes importants de l'acarien *Dermatophagoides pteronyssinus*.

D'après l'alignement multiple entre les séquences des allergènes étudiés (Der p 5, Blo t 5, Blo t 21, Der f 5, Sui m 5, Lep d 5, Der p 21 et Der f 21), nous avons pu constater que les résidus qui constituent l'épitope (appelé épitope-I dans ce chapitre) situé entre la deuxième boucle et l'hélice  $C_{terminale}$ , déterminé par Chan et *al.* (2008) et par Tan et *al.* (2012) dans Blo t 5 et Blo 21, respectivement, sont mieux conservés au sein de ces deux familles 5 et 21, par rapport aux résidus de l'épitope (appelé épitope-II dans ce chapitre) localisés dans la première boucle de Blo t 5 (Naik et *al.*, 2008).

Chan et *al.* (2008) ont montré que les mutations introduites dans l'épitope-II déterminé par Naik et *al.* (2008), ne réduisaient pas significativement la fixation de cet allergène aux IgEs. En outre, l'impact considérable de la mutation des résidus majeurs de l'épitope-I dans Blo t 5 et Blo t 21 sur la réduction de leur IgE-réactivité, montre l'importance de cet épitope-I et de l'hélice  $C_{terminale}$  dans l'allergénicité de ces allergènes des familles 5 et 21. En conséquent, et suite à notre analyse, la faible réactivité croisée observée entre les allergènes de ces deux familles 5 et 21 vient sans doute de la conservation partielle des résidus épitopes majeurs (ou aux résidus adjacents de ces derniers) se trouvant dans cet épitope-I de Blo t 5 et de Blo t 21.

D'autre part, grâce au programme SEPPA, quatre sites antigéniques ont pu être prédits dans Der p 5. Les résidus épitopes dans les parties  $N_{terminale}$  et  $C_{terminale}$ , appelés épitope- $N_{terminal}$  et

épitope-*C<sub>terminal</sub>*, respectivement, pourraient jouer un rôle dans l'allergénicité de cet allergène, soit en formant un épitope conformationnel avec d'autres résidus soit en contribuant à la réticulation des IgE liés aux récepteurs. Les deux autres épitopes ont été appelés épitope-I et épitope-II. L'épitope-I (de la position 62 à 85 de notre alignement de séquence) correspond à la région de l'épitope le plus important de Blo t 5 et de Blo t 21. L'épitope-II (de la position 35 à 46 de notre alignement de séquence) correspond à l'épitope de Blo t 5 déterminé par Naik et *al.* (2008).

Parmi les résidus qui composent ces deux épitopes, nous avons sélectionné ceux qui pourraient être importants dans la fixation aux IgEs en nous basant sur la nature des résidus (chargés, polaires et aromatiques) et leur accessibilité au solvant. Les positions ainsi sélectionnées sont: Q71, R72, K73, D74, D76, D79, Q80, Y81 et E84 dans l'épitope I, et les résidus H35, E37, E38, K39, K42, D46 et E43 dans l'épitope II. Nous avons ensuite évalué l'impact de leur mutation en alanine sur la stabilité thermodynamique de la protéine. Ces mutations n'étant pas fortement défavorables du point de vue de la stabilité thermodynamique, ont été proposées pour la validation expérimentale. Une première série de mutations est actuellement en phase de validation dans le laboratoire du "Centre de l'Ingénierie des protéines-CIP" de Liège, tandis que les autres mutations seront testées ultérieurement.

Par ailleurs, afin de mieux comprendre la variabilité de la réactivité croisée entre Der p 5 et Blo t 5, des mutations au niveau des résidus majeurs de l'épitope II de Blo t 5 ont été engendrées dans Der p 5 (K39N, P40K et A50R) et vice versa dans Blo t 5 (N39K, K40P et R50R). Cette étude nous permettra de mieux comprendre l'importance de ces résidus au niveau de cet épitope. Ces mutations sont également en cours de validation expérimentale dans le même laboratoire cité ci-dessus.

## *Chapitre III:*

*Modélisation des structures  
tridimensionnelles monomériques et  
oligomériques des allergènes d'acariens  
des familles 5 et 21*

## *Chapitre III:*

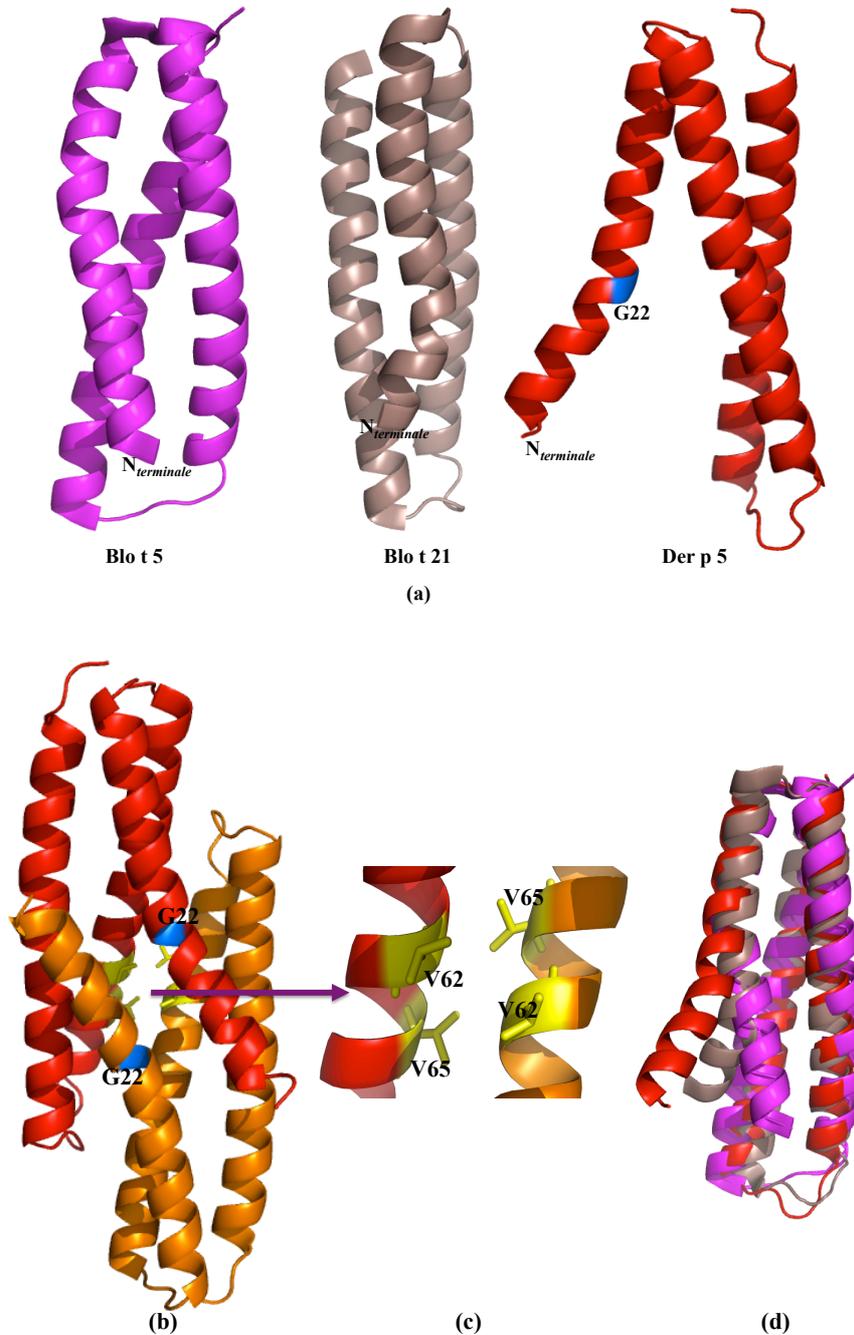
### *Introduction et problématique*

### III.1 Introduction et problématique

Une voie permettant la compréhension de la fonction biologique et des propriétés d'un allergène passe par l'étude de sa structure 3D. Parmi les allergènes des familles 5 et 21, seules les structures de Blo t 5, de Blo t 21 et de Der p 5 ont été déterminées expérimentalement par RMN ou par cristallographie aux rayons-X (Naik et *al.*, 2008; Chan et *al.*, 2008 ; Tan et *al.*, 2012; Mueller et *al.*, 2010a). En outre, la fonction biologique de ces allergènes est, à l'heure actuelle, inconnue.

Les structures obtenues par RMN de Blo t 5 (code PDB 2JMH) et de Blo t 21 (code PDB 2LM9) montrent que ces allergènes sont constitués de trois hélices alpha de taille similaire emballées dans un faisceau antiparallèle formant une organisation en *coiled-coil* (figure n°18a) (Naik et *al.*, 2008; Tan et *al.*, 2012). La structure RX de Der p 5 (code PDB 3MQ1) est similaire à celle de Blo t 5 et de Blo t 21 (figure n°18a), mais contrairement à ces dernières qui sont rapportées comme étant des monomères, la structure de Der p 5 est sous forme d'un hexamère, constitué d'un triplet d'homodimères (Mueller et *al.*, 2010a). La dimérisation de cet allergène crée une cavité hydrophobe au centre de la molécule. Une hypothèse avancée et non encore vérifiée est que cette cavité pourraient être impliquée dans la réponse immunitaire, suite à l'éventuelle liaison à un ligand (Mueller et *al.*, 2010a; Thomas, 2012).

Dans la structure dimérique de Der p 5, l'hélice N<sub>terminale</sub> est coudée au niveau de la glycine en position 22 (selon la numérotation de notre alignement multiple, voir figure n°13, chapitre II, section III.3.1) (figure n°18a). Cette G22 pourrait être importante pour la flexibilité de cette hélice et conduire à la formation du dimère où les deux chaînes de Der p 5 sont enchevêtrées (figure n° 18b). Cette glycine est également présente dans Blo t 5, mais son hélice N<sub>terminale</sub> est droite. Cette glycine est substituée par une isoleucine (I11) dans Blo t 21 mais la structure de cet allergène est coudée au niveau de son résidu phénylalanine dans la position 19 (F19). La superposition de la structure de ces trois allergènes montre la différence de la structure de leurs hélices N<sub>terminale</sub> (figure n°18d). Selon Mueller et *al.* (2010a) les différences de séquence entre Blo t 5 et Der p 5 aux positions 62 et 65 peut expliquer le fait que Der p 5 est essentiellement un dimère et Blo t 5 un monomère. Dans Der p 5, la valine 62 de la première chaîne interagit avec la valine 65 de la deuxième chaîne formant une interaction de type "valine zipper" (figure n°18c). Ces deux positions (V62 et V65) sont substituées par une alanine dans Blo t 5, et par une sérine et une tyrosine, respectivement dans Blo t 21, ce qui devrait diminuer la surface hydrophobe entre les deux chaînes et pourrait expliquer le fait que Blo t 5 et Blo t 21 restent sous forme monomérique.



**Figure n°18:** (a) Représentation de la structure 3D des allergènes Blo t 5 (code PDB 2JMH), de Blo t 21 (code PDB 2LM9) et de Der p 5 (code PDB 3MQ1). (b) Structure de dimère (chaînes A en rouge et B en orange) de Der p 5, position des G22 (en bleu) et des V62 et V65 (en jaune). (c) Zoom sur la position des interactions valine zipper (V62\_V65) entre les deux chaînes A et B de Der p 5. (d) Superposition des structures de ces trois allergènes.

D'autre part, la structure 3D de Der p 21 n'a pas été résolue expérimentalement mais a été étudiée par diffusion des rayons-X aux petits angles (*Small Angle X-ray Scattering-SAXS*) et par dichroïsme circulaire (Weghofer et al., 2008c). Les résultats de ces expériences suggèrent que cet allergène est également composé de trois hélices alpha et qu'il pourrait présenter une structure sous forme de dimère avec une forme allongée. Les simulations *ab initio* réalisées proposent un mode de dimérisation différent de celui de Der p 5.

Dans cette section, nous nous sommes intéressés à étudier la structure des allergènes des familles 5 et 21: Der f 5 et Der f 21 de *Dermatophagoides farinae*, Blo t 5 et Blo t 21 de *Blomia tropicalis*, Sui m 5 de *Suidasia medanensis*, Lep d 5 de *Lepidoglyphus destructor* et Der p 5 et Der p 21 de *Dermatophagoides pteronyssinus*, dans le but d'évaluer leur propension à dimériser et de modéliser les structures éventuelles des dimères. Cette dernière partie a pour objectif d'identifier si le processus de dimérisation de ces allergènes permettrait la création d'une cavité hydrophobe similaire à celle de Der p 5, qui pourrait être responsable de l'évolution de la réponse immunitaire de la tolérance à l'inflammation allergique au travers de la liaison à certains ligands.

Nous avons abordé ces questions par des approches bioinformatiques et des techniques de modélisation moléculaire. Dans un premier temps, nous avons générés des modèles de monomère pour les allergènes étudiés sur la base des trois structures connues de Der p 5, Blo t 5 et Blo t 21. Pour chaque allergène (Der f 5, Sui m 5, Lep d 5, Der p 21 et Der f 21), nous avons généré un premier ensemble de modèles monomériques en utilisant la structure 3D de Blo t 5 comme patron. Ces modèles présentent une structure composée de trois hélices alpha antiparallèles avec une hélice  $N_{terminale}$  droite. Nous avons par la suite généré un deuxième ensemble de modèle en nous basant sur la structure de la chaîne A de Der p 5 (3MQ1\_A). Ces modèles ont une hélice  $N_{terminale}$  coudée. Enfin, un troisième ensemble a été généré en utilisant la structure 3D de Blo t 21 comme patron. Ces modèles ont une hélice  $N_{terminale}$  légèrement coudée. Cette étape nous a permis d'estimer, selon un critère énergétique, la propension que présente les allergènes étudiés (Der f 5, Sui m 5, Lep d 5, Der p 21 et Der f 21) à adopter une structure avec une hélice  $N_{terminale}$  droite ou coudée. Par la suite, les modèles générés sur la base de la structure de Der p 5 et retenus selon un critère énergétique ont été utilisés pour effectuer des simulations d'amarrage protéine-protéine afin de prédire leur mode de dimérisation.

Nous tenons à souligner que les logiciels utilisés pour prédire les structures des dimères possibles de ces protéines ne précisent pas si le dimère pourrait effectivement se former ou non, mais permettent d'évaluer si la structure de l'éventuel dimère pourrait être similaire à celle de Der p 5.

*Chapitre III:*

*Méthodes*

## III.2 Méthodes

### III.2.1 Modélisation comparative des structures 3D des allergènes des familles 5 et 21

La première étape de ce travail consiste en la génération des modèles de structure 3D des allergènes des familles 5 et 21 (Der f 5, Sui m 5, Lep d 5, Der f 21 et Der p 21) en se basant sur les structures de Blo t 5, de Blo t 21 et de Der p 5 résolues expérimentalement.

Nous avons généré la structure de nos allergènes par la méthode de modélisation comparative en utilisant le programme "MODELLER" (Sali & Blundell, 1993). Ce programme a besoin d'un alignement de séquences entre la séquence de la protéine à modéliser (la cible) et la séquence d'une protéine similaire dont la structure est connue (le patron, dans notre cas la structure de Blo t 5, de Blo t 21 ou celle de Der p 5). Par la suite, des modèles de structure de la cible sont générés par satisfaction des contraintes spatiales. Dans notre travail, nous avons extrait les alignements par paire entre cible et patron à partir de l'alignement multiple de nos allergènes étudiés (figure n°13, chapitre II, section II.3.1).

#### III.2.1.1 Création des modèles

Nous avons utilisé le programme MODELLER 9v4 pour établir un modèle de la structure 3D monomérique de Der f 5, Sui m 5, Lep d 5, Der p 21 et Der f 21 en utilisant les structures expérimentales de Blo t 5 (code pdb 2JMH), de Blo t 21 (code pdb 2LM9) et de Der p 5 (code pdb 3MQ1) comme patrons.

A partir de l'alignement multiple, nous avons extrait des alignements par paire entre les patrons (Blo t 5, Blo t 21 et Der p 5) et les séquences cibles (Der f 5, Sui m 5, Lep d 5, Der p 21 et Der f 21), respectivement, que nous avons introduit comme entrée pour MODELLER.

Trois types de modèles de monomères ont été modélisés par MODELLER. Un premier ensemble de modèles a été généré pour les allergènes Der f 5, Sui m 5, Lep d 5, Der p 21 et Der f 21 en utilisant la structure de Blo t 5 comme patron. Ces modèles présentent une hélice alpha  $N_{terminale}$  droite. Le deuxième et le troisième ensemble de modèles de ces allergènes ont été générés en se basant sur la structure de la chaîne A de Der p 5 (3MQ1\_A) et de la structure de Blo t 21 (2LM9), respectivement. Ces modèles ont une hélice  $N_{terminale}$  coudée, le degré de courbure observé au sein de 2LM9 étant moindre que celui dans 3MQ1. Par ailleurs, des modèles avec une hélice  $N_{terminale}$  coudée comme celle de Der p 5 ont également été

construits pour Blo t 5 et pour Blo t 21 afin d'étudier l'impact de la flexibilité de cette hélice dans le mode de dimérisation de ces deux allergènes. Trente modèles ont été générés pour chaque protéine avec chacun des trois patrons en utilisant les paramètres standard de MODELLER.

### III.2.1.2 Evaluation de la qualité et sélection des modèles 3D

Evaluer la qualité des structures 3D est une étape cruciale dans la construction de modèles. Dans cette étude, l'évaluation de la qualité des modèles 3D prédits a été réalisée à l'aide de deux types de fonctions d'énergie: une semi-empirique (Gromos, Van Der Spoel et *al.*, 2005) pour la minimisation de l'énergie de la structure des modèles et une dérivée à partir de bases de données de structure connues (Anolea, Melo & Feytmans, 1998). En effet, la minimisation d'énergie a pour but de ramener la fonction énergétique globale à un minimum que l'on espère le plus proche de l'énergie de la structure native en éliminant d'éventuels contacts inappropriés. Cette minimisation a été réalisée dans le vide, à l'aide du champ de force GROMOS96 45a3 avec le programme « Gromacs » (*GROningen MACHine for Chemical Simulations*) (Van Der Spoel et *al.*, 2005), sur chacun des modèles proposé par MODELLER. L'algorithme de minimisation utilisé dans ce programme est de type "de gradients de plus grande pente" (« *Steepest Decent* »). En outre, le score Anolea (Melo & Feytmans, 1998) a également été calculé pour chaque modèle après minimisation de l'énergie avec Gromos. Anolea calcul un score d'énergie en évaluant l'environnement spatial pour chaque atome dans le modèle en comparaison avec l'environnement observé dans des structures résolues par cristallographie aux rayons-X à haute résolution.

Ces deux scores (Gromos et Anolea) étant calculés différemment, nous les avons normalisés en Z-score selon l'équation suivante :

$$Z - score_i = \frac{S_i - \langle S \rangle}{\sigma} \quad (1)$$

où  $S_i$  est le score obtenu pour le modèle considéré  $i$ ,  $\langle S \rangle$  est la moyenne des scores de tous les modèles et  $\sigma$  est l'écart type de la distribution des scores.

Cette normalisation permet de comparer les 30 modèles générés pour chaque allergène avec chacun des patrons selon les deux scores énergétiques, Gromos et Anolea. Le Z-score final pour chaque modèle est la moyenne des Z-scores calculés pour chacune de ces deux fonctions d'énergie. Nous avons défini le meilleur modèle comme étant celui de plus petit Z-score

moyen. Ainsi, les modèles ayant un Z-score moyen inférieur ou égal à -0,5 ont été sélectionnés.

### III.2.2 Homodimérisation

#### III.2.2.1 Simulations d'amarrage protéine-protéine

Les modèles de monomères retenus selon notre critère énergétique et générés sur la base de la chaîne A de Der p 5 (code PDB 3MQ1) pour Der f 5, Blo t 5, Sui m 5, Lep d 5, Der p 21, Der f 21 et Blo t 21 ont été soumis au programme ClusPro 2.0 (<http://cluspro.bu.edu/>; Kozakov et al., 2006) afin de générer les homodimères éventuels pour chaque allergène.

ClusPro est l'un des programmes d'amarrage les plus performant au tour 13-19 de CAPRI (*Critical Assessment of Predicted Interactions*). Sa fonction de score prend en compte l'énergie libre de désolvation et les interactions électrostatiques. Il sélectionne les 1000 meilleurs scores et les regroupe dans des clusters (amas). Chaque cluster est caractérisé par son nombre de solutions, le score ClusPro du centre du cluster et le score ClusPro le plus bas trouvé dans le cluster. Pour chaque monomère, nous avons retenu le dimère ayant le plus bas score ClusPro.

#### III.2.2.2 Calcul de l'écart quadratique moyen (RMSD)

Afin d'établir les similarités structurales entre les différentes conformations, le programme U3BEST (Kabsch, 1978) a été utilisé pour calculer le RMSD (*Root Mean Square Deviation*; écart quadratique moyen) (Equation 2). Il s'agit de la distance moyenne minimale entre atomes correspondant après superposition des structures.

$$RMSD = \min\left(\sqrt{\frac{1}{N} \sum_i^N [(x_{1,i} - x_{2,i})^2 + (y_{1,i} - y_{2,i})^2 + (z_{1,i} - z_{2,i})^2]}\right) \quad (2)$$

Où,

$(x_1, y_1, z_1)$  et  $(x_2, y_2, z_2)$  sont les coordonnées cartésiennes des atomes correspondants dans les protéines 1 et 2 et N est le nombre d'atomes superposés.

Plus le RMSD est bas, plus les structures 1 et 2 sont proches.

### III.2.2.3 Assignation des interactions à l'interface des dimères

Le programme PIC "*Protein Interactions Calculator*" (Tina et al., 2007) a été utilisé pour calculer les interactions d'interface trouvées dans les dimères. Le pourcentage des interactions de type  $i$  (ponts salins, contacts hydrophobes et les liaisons hydrogènes) trouvées à l'interface des dimères X et Y,  $P_{X/Y}^i$  a été calculé en se basant sur l'équation suivante:

$$P_{X/Y}^i = \frac{N_{X/Y}^i}{N_Y^i} \cdot 100\% \quad (3)$$

Où,

$N_{X/Y}^i$  est le nombre d'interactions de type  $i$  communes à l'interface des dimères X et Y,

$N_Y^i$  est le nombre des interactions de type  $i$  à l'interface du dimère Y.

### III.2.2.4 Identification des cavités

Le programme CASTp "*Computed Atlas of Surface Topography of Proteins with structural and topographical mapping of functionally annotated residues*" (<http://cast.engr.uic.edu>; Dundas et al., 2006) a été utilisé afin d'identifier les cavités au sein des modèles 3D générés pour les allergènes étudiés.

## *Chapitre III:*

### *Résultats et discussions*

### **III.3 Résultats et discussions**

Le but de cette partie de recherche a été d'étudier le mode de dimérisation des allergènes des familles 5 et 21 (Der f 5, Blo t 5, Sui m 5, Lep d 5, Der f 21, Der p 21 et Blo t 21). Ainsi, nous avons utilisé des approches de modélisation moléculaire pour étudier la propension de ces allergènes à adopter des homodimères similaires à celui de Der p 5 (code PDB 3MQ1) avec la création d'une grande cavité hydrophobe qui pourrait accueillir des ligands. Cette étude nous permet d'éclairer nos connaissances quant à la fonction probable de ces allergènes des familles 5 et 21.

#### **III.3.1 Analyse des résidus de l'interface de Der p 5 dans l'alignement multiple de séquence**

L'alignement multiple de nos séquences réalisé par ClustalW (voir chapitre II, section **II.2.1.1**) reflète le degré de conservation de résidus entre ces séquences (figure n°19). Les résidus qui interviennent dans les interactions à l'interface de la structure expérimentale du dimère de Der p 5 (chaînes A et B du code PDB 3MQ1, 3MQ1\_AB) ont été identifiés par le programme PIC et sont présentés en jaune dans la figure n°19. Ces interactions (contacts hydrophobes, liaisons hydrogènes et ponts salins) entre les deux chaînes de Der p 5 jouent un rôle important dans la formation et la stabilité du dimère de cet allergène. Les pourcentages d'identité et de similarité par rapport à ces résidus d'interface entre Der p 5 et les allergènes étudiés (Der f 5, Blo t 5, Sui m 5, Lep d 5, Der f 21, Der p 21 et Blo t 21) sont donnés dans le tableau n°10. Les résultats montrent que ces pourcentages sont plus grands que ceux calculés sur l'ensemble des séquences (tableau n°8, chapitre II, section **II.3.1**), sauf pour Lep d 5. Nous remarquons aussi une très forte similarité de séquence (95%) entre Der p 5 et Der f 5 sur ces résidus de l'interface (tableau n°10).

N° dans le PDB 3MQ1	34										40					45														
N° de l'alignement	1					5					10					15					20					25				
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 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					
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 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					
Blo t 21	D	N	F	R	H	E	F	D	H	M	I	V	N	T	A	T	Q	R	F	H	E	I	E	K	F					
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					

N° dans le PDB 3MQ1	50					55					60					65					70				
N° de l'alignement	30					35					40					45					50				
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 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
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	M	K	E	K	I	V	R
Lep d 5	L	A	H	L	S	H	E	V	T	E	L	E	K	S	K	S	K	E	L	K	A	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
Blo t 21	L	L	H	I	T	H	E	V	D	E	L	E	K	T	G	N	K	D	E	K	A	R	L	L	R
Der p 21	L	L	H	L	S	E	E	Y	K	E	L	E	K	T	K	S	K	E	L	K	E	Q	I	L	R

N° dans le PDB 3MQ1	75					80					85					90					95				
N° de l'alignement	55					60					65					70					75				
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 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
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 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
Blo t 21	E	L	T	V	S	E	A	F	I	E	G	S	R	G	Y	F	Q	R	E	L	K	R	T	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

N° dans le PDB 3MQ1	100					105					110					115					120				
N° de l'alignement	80					85					90					95					100				
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 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
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 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
Blo t 21	D	L	L	E	K	F	N	F	E	A	A	L	A	T	G	D	L	L	L	K	D	L	K	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

N° dans le PDB 3MQ1	125																								
N° de l'alignement	105					110																			
Blo t 5	E	Q	K	V	K	D	I	Q	T	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Der p 5	E	A	R	V	K	K	I	E	V	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Der f 5	E	Q	R	V	K	K	I	E	V	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sui m 5	E	A	K	V	K	A	I	N	V	H	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Lep d 5	A	T	K	V	K	A	I	E	T	K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Der f 21	A	K	K	V	K	A	V	K	S	D	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Blo t 21	Q	K	R	V	Q	D	S	E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Der p 21	A	K	K	V	K	A	V	N	P	D	E	Y	Y	-	-	-	-	-	-	-	-	-	-	-	-

**Figure n°19:** Alignement multiple des séquences des allergènes étudiées obtenu par ClustalW. La première ligne contient la numérotation des résidus dans la structure cristallographique de Der p 5 (3MQ1). La ligne suivante correspond à la numérotation de l'alignement de séquences. Les résidus intervenant dans les interactions à l'interface du dimère de Der p 5 (chaînes A et B de 3MQ1) sont colorés en jaune. Ces résidus d'interface de Der p 5 conservés dans les autres allergènes sont colorés en jaune également. Ceux pour lesquels le caractère hydrophobe, le caractère acide ou le caractère polaire est conservé sont en vert, rose et bleu, respectivement (Khemili et al., 2012).

**Tableau n°10:** Pourcentages d'identité et de similarité de séquences entre Der p 5 et les autres allergènes, calculés au niveau des résidus impliqués dans les interactions à l'interface du dimère 3MQ1\_AB. La similarité est définie comme étant le pourcentage d'acides aminés pour lesquels le score d'alignement de séquence (selon notre matrice de score, Blossum62) est positif.

	Identité	Similarité
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%
Blo t 21	33%	75%

Les allergènes Lep d 5, Der p 21, Der f 21 et Blo t 21 ne contiennent pas de résidu glycine en position 22 de l'alignement (figure n°19) et montrent un faible pourcentage d'identité pour les résidus d'interface de dimère de Der p 5. D'après Mueller et *al.* (2010a), cette glycine devrait jouer un rôle important dans la formation d'un coude dans l'hélice  $N_{terminale}$  et d'une structure dimérique enchevêtrée de cet allergène. L'absence de ce résidu dans ces quatre allergènes (Lep d 5, Der p 21, Der f 21 et Blo t 21) nous laisse penser que ces allergènes n'adopteraient pas une structure dimérique similaire à Der p 5.

Par ailleurs, selon Mueller et *al.* (2010a), les valines V62 et V65 (figure n°19) jouent un rôle dans la stabilité du dimère de Der p 5 en formant une interaction de type "valine zipper". Ces deux valines sont conservées uniquement dans Der f 5.

Contrairement à ce qu'on a trouvé pour Lep d 5, Der p 21, Der f 21 et de Blo t 21, le calcul des pourcentages d'identité et de similarité de séquences ainsi que la conservation des résidus de l'interface entre Der p 5 et les autres allergènes nous laisse suggérer que Blo t 5, Der f 5 et Sui m 5 pourraient adopter une structure dimérique semblable à celle de Der p 5 avec une grande cavité entre les deux monomères. La conservation de séquence la plus grande a été trouvée entre Der p 5 et Der f 5.

L'étape suivante de notre travail est la génération des modèles de structure des allergènes étudiés (Der f 5, Sui m 5, Lep d 5, Der f 21 et Der p 21). La similarité de séquence entre ces allergènes avec leurs patrons (Blo t 5, Blo t 21 et Der p 5) étant suffisamment élevée (tableau n°8, chapitre II, section **II.3.1**), nous avons utilisé une approche de modélisation comparative pour prédire la structure des monomères des allergènes et ensuite, des modèles de leurs structures dimériques ont été conçus par la méthode d'amarrage protéine-protéine, en ne considérant pas de flexibilité.

### III.3.2 Modélisation comparative de la structure des monomères

Nous avons utilisé les structures 3D de Blo t 5 (code PDB 2JMH), de Blo t 21 (code PDB 2LM9) et de la chaîne A de Der p 5 (code PDB 3MQ1\_A) comme patrons pour créer un modèle de la structure 3D de Der f 5, Sui m 5, Lep d 5, Der f 21 et Der p 21 en utilisant MODELLER 9v4 (les scripts utilisés pour la modélisation sont donnés dans l'annexe I). Ces patrons présentent des différences au niveau de la flexibilité de leur hélice  $N_{terminale}$  qui est droite dans Blo t 5, légèrement coudée dans Blo t 21 et coudée dans Der p 5 (figure n°18). Cependant, connaissant le rôle de la flexibilité de cette hélice dans la formation du dimère de Der p 5, nous avons également généré un modèle présentant une hélice  $N_{terminale}$  coudée pour Blo t 5 et pour Blo t 21, et ce en utilisant la chaîne A de 3MQ1 en tant que patron.

Un premier ensemble de modèles a été obtenu avec la structure de Blo t 5 (2JMH) en tant que patron. Ces modèles présentent une hélice  $N_{terminale}$  droite. L'énergie de ces modèles a été évaluée avec Anolea et Gromos (voir Méthodes, section **III.2.1.2**). Les valeurs d'énergie des modèles sélectionnés selon notre critère d'énergie (Z-score moyen inférieur ou égale à -0,5, voir Méthodes, sections **III.2.1.2**) sont données dans le tableau A1 (voir annexe II). Le nombre de modèles retenus pour ce patron est de: 10 pour Der f 5 (derf5\_monomer1→10[2JMH]), 10 pour Sui m 5 (suim5\_monomer1→10[2JMH]), 9 pour Lep d 5 (lepd5\_monomer1→9[2JMH]), 8 pour Der f 21 (derf21\_monomer1→8[2JMH]) et 9 pour Der p 21 (derp21\_monomer1→9[2JMH]).

Un deuxième ensemble de modèles a été créé sur la base de la structure de Blo t 21 (2LM9) comme patron. L'hélice  $N_{terminale}$  de ces modèles est légèrement coudée. L'énergie de ces modèles est donnée dans le tableau A2 (voir annexe II). Le nombre de modèles générés que nous avons retenus est de: 9 pour Der f 5 (derf5\_monomer1→9[2LM9]), pour Sui m 5 (suim5\_monomer1→5[2LM9]), 7 pour Lep d 5 (lepd5\_monomer1→7[2LM9]), 8 pour

Der f 21 (derf21\_monomer1→8[2LM9]) et 9 pour Der p 21 (derp21\_monomer1→9[2LM9]).

Le dernier ensemble de modèles de ces allergènes a été calculé en utilisant la structure de la chaîne A de Der p 5 (3MQ1\_A) comme patron. L'hélice  $N_{terminale}$  est coudée dans ces modèles. L'énergie de ces modèles est donnée dans le tableau A3 (voir annexe II). Le nombre de modèles générés en se basant sur la structure de la chaîne A de Der p 5 que nous avons retenus est de: 5 pour Der f 5 (derf5\_monomer1→5[3MQ1\_A]), 7 pour Sui m 5 (suim5\_monomer1→7[3MQ1\_A]), 6 pour Lep d 5 (lepd5\_monomer1→6[3MQ1\_A]), 7 pour Der f 21 (derf21\_monomer1→7[3MQ1\_A]), 8 pour Der p 21 (derp21\_monomer1→8[3MQ1\_A]), 8 pour Blo t 5 (blot5\_monomer1→8[3MQ1\_A]) et 7 pour Blo t 21 (blot21\_monomer1→7[3MQ1\_A]).

Les valeurs de plus basse énergie obtenues avec chaque modèle et l'énergie moyenne calculée avec Gromos et Anolea des modèles sélectionnés sont données dans le tableau n°11. La comparaison de ces valeurs indique, pour chaque allergène, si la structure avec une hélice  $N_{terminale}$  droite (patron 2JMH) est prédite avec un meilleur score que celle avec une hélice légèrement coudée (patron 2LM9) ou une hélice coudée (patron 3MQ1\_A). Pour Der f 5, la structure modélisée avec le patron 3MQ1\_A est clairement préférée par rapport à celle obtenue avec le patron 2JMH et avec 2LM9, et ce quel que soit le critère utilisé. Les résultats sont plus contrastés pour les modèles de Sui de 5 m, Lep d 5, Der p 21 et Der f 21. Nous arrivons à des conclusions différentes selon le critère utilisé pour définir quel patron conduit au meilleur modèle (énergie la plus faible ou énergie moyenne calculée avec Anolea ou Gromos). Ces résultats suggèrent qu'une hélice  $N_{terminale}$  coudée serait plus probable dans le monomère de Der f 5 que dans les autres allergènes étudiés. Notons que les séquences de Lep d 5, Der p 21 et Der f 21 ne présentent pas un résidu glycine en position 22 de l'alignement. La présence d'un autre résidu à cette position peut probablement diminuer la propension à former un coude dans l'hélice  $N_{terminale}$  de ces allergènes.

**Tableau n°11:** Valeurs d'énergie, calculées avec Anolea et Gromos, des modèles monomériques générés sur la base des patrons 2JMH, 2LM9 et 3MQ1\_A

	La plus basse énergie			La plus basse énergie			Energie moyenne			Energie moyenne		
	Anolea			Gromos			Anolea			Gromos		
Patron	3MQ1_A	2LM9	2JMH	3MQ1_A	2LM9	2JMH	3MQ1_A	2LM9	2JMH	3MQ1_A	2LM9	2JMH
Der f 5	<b>-232,0</b>	-160,0	-190,0	<b>-11283,3</b>	-9360,6	-9805,5	<b>-207,2</b>	-138,4	-173,3	<b>-10008,0</b>	-7372,6	-8427,9
Sui m 5	-186,0	<b>-204,0</b>	-111,0	-8609,9	-7549,0	<b>-9078,4</b>	-151,0	<b>-155,3</b>	-79,3	<b>-8036,2</b>	-7122,7	-7662,0
Lep d 5	<b>-232,0</b>	-192,0	<b>-232,0</b>	<b>-8569,8</b>	-6939,3	-8320,8	<b>-194,3</b>	-177,6	-186,4	-7452,4	-5928,3	<b>-7748,1</b>
Der p 21	<b>-189,0</b>	-182,0	-109,0	-8627,2	-8819,8	<b>-8963,9</b>	<b>-144,8</b>	-144,3	-72,2	<b>-8075,4</b>	-6620,6	-7778,1
Der f 21	-188,0	<b>-206,0</b>	-173,0	-7215,5	-7237,4	<b>-7297,1</b>	<b>-133,0</b>	-168,1	-101,0	<b>-6731,0</b>	-5917,1	-5790,9
Blo t 5	-235,0	/	/	-10256,0	/	/	-154,9	/	/	-9188,5	/	/
Blo t 21	-241,0	/	/	-7770,1	/	/	-206,9	/	/	-7013,6	/	/

Nous avons également utilisé le patron 3MQ1\_A pour obtenir une structure coudee pour les monomères de Blo t 5 et de Blo t 21. Ce qui est intéressant, c'est que l'énergie d'une telle structure pour Blo t 5 est assez basse, pas autant que l'énergie obtenue pour le modèle de Der f 5, mais inférieure à celles calculées pour les modèles de monomères de Sui 5 m, Lep d 5, Der p 21 et Der f 21 (tableau n°11). Ce résultat est également observé pour Blo t 21 mais uniquement sur les scores d'énergie d'Anolea.

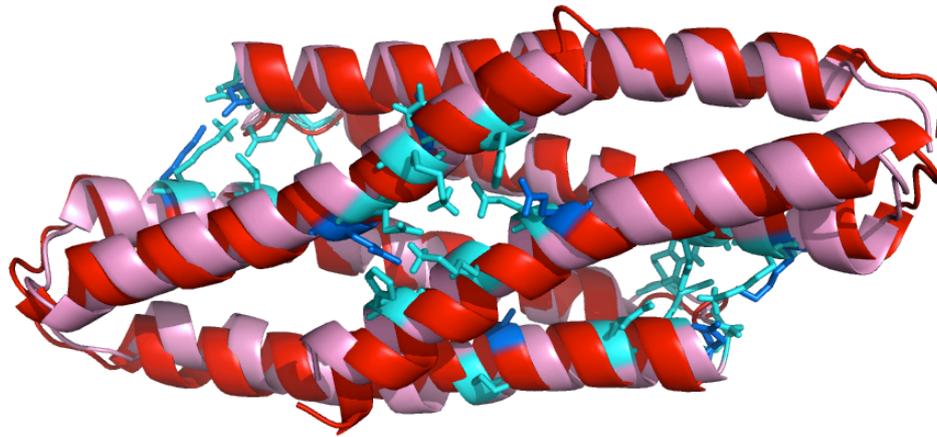
### III.3.3 Homodimérisation

#### III.3.3.1 Fiabilité de ClusPro pour prédire la structure des dimères

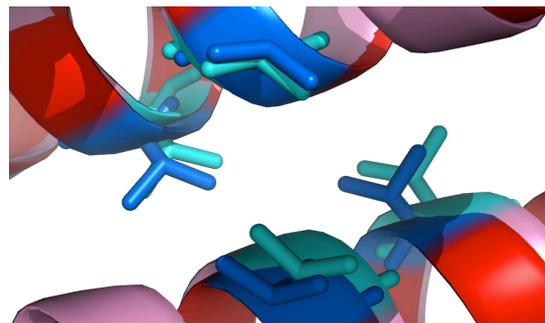
Nous avons utilisé le serveur ClusPro2.0 pour prédire la structure des dimères. Pour tester la capacité de ClusPro à prédire la structure dimérique de ces allergènes, nous avons d'abord utilisé ce programme pour prédire la structure du dimère de Der p 5 et ensuite comparé cette structure prédite avec la structure expérimentale composée de ses deux chaînes A et B (3MQ1\_AB).

La chaîne A de Der p 5 (3MQ1\_A) a été soumise au programme ClusPro et 20 solutions ont été générées en sortie. La structure ayant le plus bas score ClusPro, appelée derp5\_dimer[3MQ1\_A], a été identifiée comme celle qui est la plus proche de la structure expérimentale de Der p 5 (3MQ1\_AB). L'évaluation de la similarité entre cette structure prédite (derp5\_dimer[3MQ1\_A]) et la structure expérimentale (3MQ1\_AB) de Der p 5 s'est basée sur (i) le calcul du rmsd entre ces structures après superposition et sur (ii) les interactions présentes à l'interface de ces deux structures de dimères. La valeur de rmsd après superposition des carbones alpha ( $C\alpha$ ) entre les deux structures 3MQ1\_AB et derp5\_dimer[3MQ1\_A] est égale à 1,9 Å (figure n°20a). Ce résultat reflète la forte similarité structurale entre la structure prédite et la structure cristallographique du dimère de Der p 5.

Les interactions à l'interface de 3MQ1\_AB et de derp5\_dimer [3MQ1\_A] ont été identifiées par le programme PIC. La comparaison de ces interactions montre que (figure n°21): 36 interactions ont été trouvées dans la structure expérimentale de Der p 5 3MQ1\_AB (24 contacts entre les résidus hydrophobes, 5 liaisons hydrogène et 7 ponts salins) et 57 interactions ont été trouvées dans la structure prédite derp5\_dimer[3MQ1\_A] (33 contacts entre les résidus hydrophobes, 14 liaisons hydrogène et 10 ponts salins).

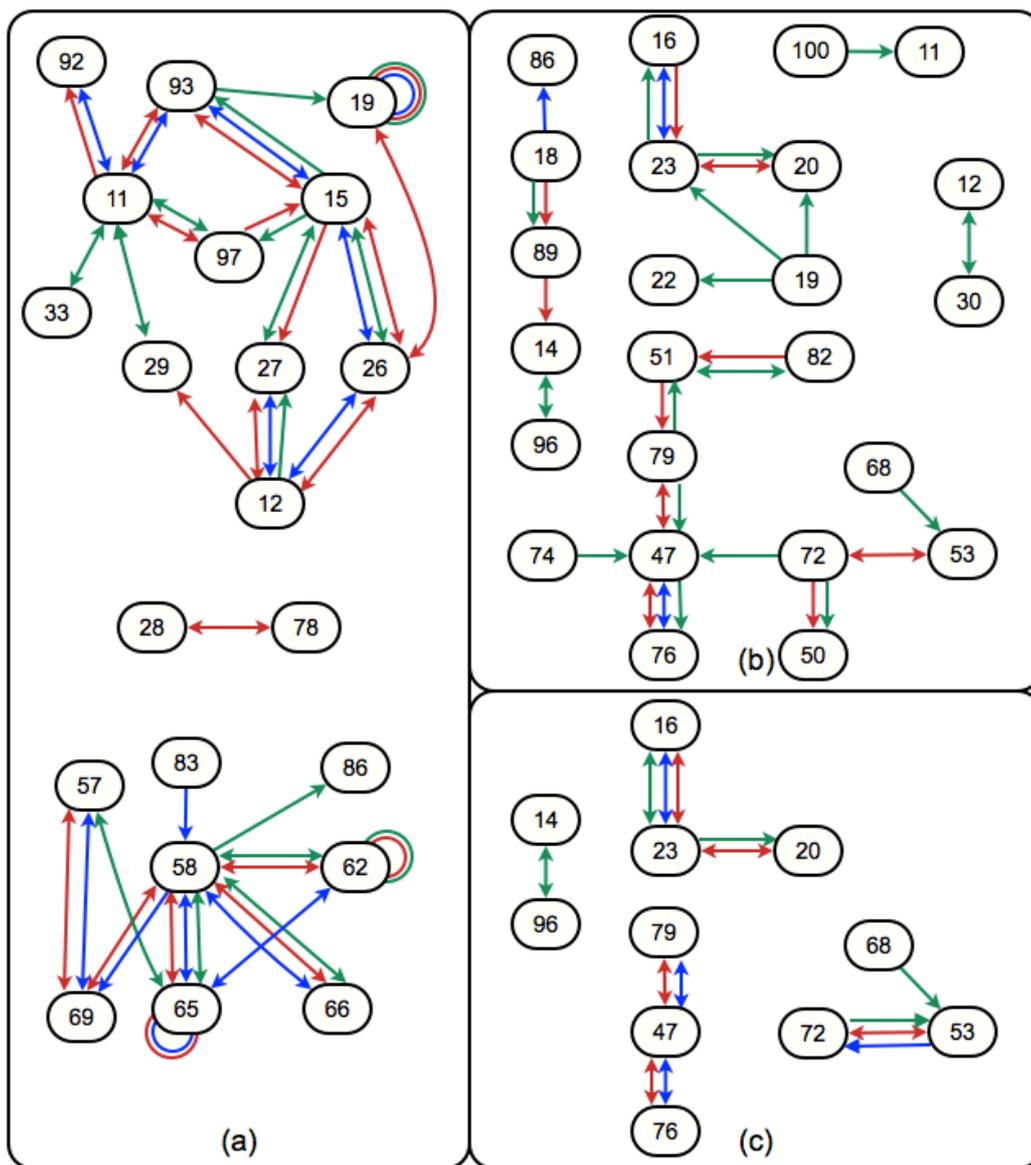


(a)



(b)

**Figure n°20:** Superposition entre la structure dimérique expérimentale de Der p 5 '3MQ1\_AB'(chaines A et B du code pdb 3MQ1 représentée en bleu clair) et la structure du dimère de Der p 5 prédit par ClusPro 'derp5\_dimer[3MQ1\_A]' (en saumon). (a) Les résidus impliqués dans des liaisons hydrogènes au niveau de l'interface de 3MQ1\_AB et de Derp5\_dimer[3MQ1\_A] sont représentés en bâtonnets bleu et cyan, respectivement. (b) les valines en positions 62 et 65 (V62 et V65) qui interagissent dans l'interface de 3MQ1\_AB sont représentées en bleu et celles trouvées dans Derp5\_dimer[3MQ1\_A] sont en cyan (Khemili et *al.*, 2012).



**Figure n°21:** Interactions de type (a) contacts hydrophobes, (b) liaisons hydrogène et (c) ponts salins calculées par PIC entre les résidus de l'interface des dimères 3MQ1\_AB, derp5\_dimer [3MQ1\_A] et derf5\_dimer5. Les positions des résidus sont numérotées en fonction de la numérotation de l'alignement de séquences. Les interactions à l'interface trouvées dans la structure expérimentale de Der p 5 (3MQ1\_AB), dans la structure dimère de Der p 5 prédite par ClusPro (derp5\_dimer [3MQ1\_A]) et dans la structure du dimère prédite de Der f 5 (derf5\_dimer5) sont représentées par une flèche bleue, rouge et verte, respectivement. Les dimères ne sont pas parfaitement symétriques et les interactions détectées entre les chaînes A et B ne sont pas systématiquement les mêmes que celles entre les chaînes B et A. Les flèches représentent les interactions entre les résidus des deux chaînes et sont orientées de la chaîne A vers la chaîne B (Khemili et al., 2012).

Le tableau n°12 indique que 83% des interactions de l'interface sont trouvées en commun dans 3MQ1\_AB et derp5\_dimer[3MQ1\_A], et que 53% des interactions de l'interface trouvées dans derp5\_dimer[3MQ1\_A] sont communes avec 3MQ1\_AB.

**Tableau n°12:** Pourcentage des interactions communes à l'interface de 3MQ1\_AB, de derp5\_dimer[3MQ1\_A] et de derf5\_dimers

	3MQ1_AB			derp5_dimer[3MQ1_A]			derf5_dimers		
	$P_{3MQ1/X}^{hydrophobe}$	$P_{3MQ1/X}^{l-hydrogène}$	$P_{3MQ1/X}^{pont-salin}$	$P_{derp5/X}^{hydrophobe}$	$P_{derp5/X}^{l-hydrogène}$	$P_{derp5/X}^{pont-salin}$	$P_{derf5/X}^{hydrophobe}$	$P_{derf5/X}^{l-hydrogène}$	$P_{derf5/X}^{pont-salin}$
3MQ1_AB	/			83%	60%	100%	38%	40%	43%
				$P_{derp5/X} = 83\%$			$P_{derf5/X} = 39\%$		
derp5_dimer	61%	21%	70%	/			45%	43%	40%
[3MQ1_A]	$P_{3MQ1/X} = 53\%$						$P_{derf5/X} = 44\%$		
derf5_dimers	36%	10%	43%	60%	30%	57%	/		

Ces pourcentages sont plus élevés quand ils sont calculés seulement sur les contacts hydrophobes ou sur les ponts salins, et plus bas sur les liaisons hydrogènes. Il y a plus de liaisons hydrogènes à l'interface de derp5\_dimère[3MQ1\_A] que dans la structure expérimentale, mais la plupart des liaisons hydrogène supplémentaires sont situées dans la même région que celles dans la structure native (figure n°20a).

Par ailleurs, Mueller et *al.* (2010a) ont suggéré que les valines aux positions 62 et 65 (numérotation de l'alignement de séquences) pourraient jouer un rôle important dans la stabilité de la structure dimérique de Der p 5 et pourraient être responsables de la différence dans la propension de dimérisation entre Der p 5 et Blo t 5. Ces contacts hydrophobes ne sont pas identifiés par le programme PIC dans derp5\_dimer[3MQ1\_A] (figure n°21). Toutefois, l'analyse de la structure montre que ces deux résidus V62 et V65 présentent la même orientation dans 3MQ1\_AB et derp5\_dimer [3MQ1\_A] (figure n°20b).

Nous avons ensuite caractérisé la cavité qui pourrait être impliquée dans la fixation du ligand dans 3MQ1\_AB et dans derp5\_dimer[3MQ1\_A] avec le programme CASTp. Le résultat montre que la grande cavité qui se trouve à l'interface des deux monomères présente un volume de  $3051\text{\AA}^3$  dans 3MQ1\_AB et de  $2914\text{\AA}^3$  dans derp5\_dimer[3MQ1\_A]. En outre, 83% des résidus impliqués dans la cavité de 3MQ1\_AB sont également présents dans la cavité de derp5\_dimer[3MQ1\_A].

Ces résultats suggèrent que ClusPro est en mesure de prédire de façon fiable la structure des dimères des allergènes étudiés dans ce travail et que le modèle avec le score le plus bas de ClusPro correspond à la prédiction avec le plus bas rmsd par rapport à la structure native.

### III.3.3.2 Prédiction de la structure dimérique des allergènes et analyse de leurs propriétés

Pour tester si les différents allergènes étudiés (Der f 5, Sui m 5, Lep d 5, Der f 21, Der p 21, Blo t 5, et Blo t 21) pourraient adopter une structure dimérique possédant une cavité hydrophobe similaire à celle de Der p 5, nous avons soumis à ClusPro tous les modèles de monomères sélectionnés qui ont une hélice  $N_{terminale}$  coudée (voir le tableau A3 en annexe II). Ces monomères ont été modélisés en utilisant 3MQ1\_A comme patron. Nous avons obtenu plusieurs solutions possibles avec ClusPro, et la solution dont le score ClusPro est le plus bas a été sélectionnée. En effet, les résultats obtenus dans la section précédente suggèrent que le

score ClusPro est capable de discriminer nos présents modèles et que la structure avec le score ClusPro le plus bas devrait être la plus semblable à la structure native.

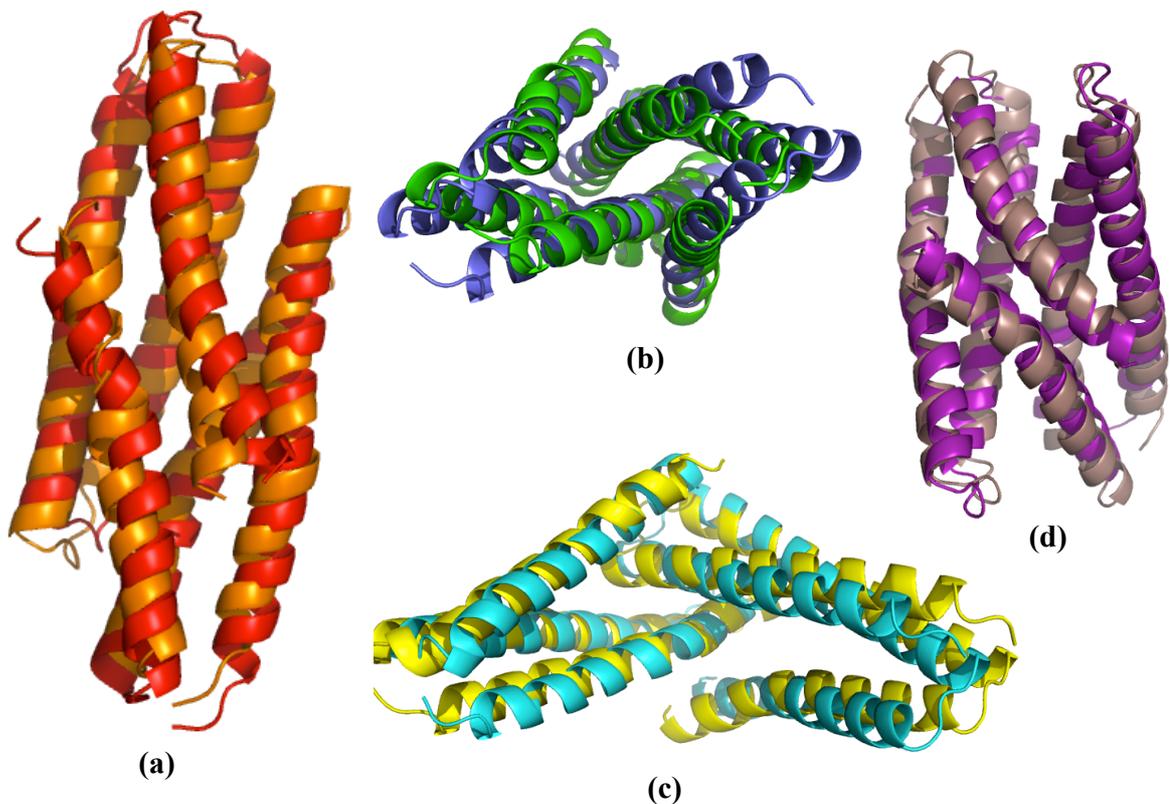
Nous avons finalement obtenu cinq homodimères pour Der f 5 (derf5\_dimer1→5), sept pour Sui 5 m (suim5\_dimer1→7), six pour Lep d 5 (lepd5\_dimer1→6), huit pour Der p 21 (derp21\_dimer1→8), sept pour Der f 21 (derf21\_dimer1→7), huit pour Blo t 5 (blot5\_dimer1→8) et sept pour Blo t 21 (blot21\_dimer1→7). La comparaison entre les dimères sélectionnés avec la structure cristallographique du dimère de Der p 5 a été effectuée par la mesure du rmsd et le calcul des pourcentages de conservation des interactions à l'interface de ces dimères. Les scores ClusPro des meilleures solutions homodimériques ainsi que leur rmsd avec 3MQ1\_AB sont présentés dans le tableau n°13 (les valeurs de scores et des rmsd de tous les homodimères prédits sont résumés dans le tableau A4 en Annexe II).

Der f 5 est le seul allergène qui présente une structure dimérique prédite similaire à celle de Der p 5 (tableau n°12, derf5\_dimer5 et figure n°22). La valeur du rmsd après superposition entre Der p 5 (3MQ1\_AB) et derf5\_dimer<sub>5</sub> est égale à 3,0 Å et 48% des interactions à l'interface identifiées dans derf5\_dimer<sub>5</sub> sont communes à celles trouvées dans derp5\_dimer[3MQ1\_A] (figure n°21 et tableau n°12). Nous avons montré précédemment que de légères différences dans l'orientation des chaînes latérales peuvent influencer sur les interactions à l'interface détectées par PIC, conduisant à un plus grand nombre d'interactions identifiées dans derp5\_dimer[3MQ1\_A] que dans 3MQ1\_AB. Pour cette raison, nous avons estimé qu'il était plus judicieux de comparer les interactions de l'interface de derf5\_dimer<sub>5</sub> à celles trouvées dans derp5\_dimère[3MQ1\_A]. Le volume de la cavité obtenue après assemblage de deux monomères dans derf5\_dimer<sub>5</sub> est égal à 1979 Å<sup>3</sup>.

**Tableau n°13** : Valeurs de la fonction de score de ClusPro des meilleurs dimères prédits pour les allergènes étudiés en utilisant 3MQ1\_A comme patron et valeurs du rmsd obtenues après superpositions de ces modèles avec la structure cristallographique 3MQ1\_AB (le dimère numéro 6, par exemple, correspond au meilleur modèle prédit par ClusPro obtenu avec le monomère numéro 6)

	Fonction de score ClusPro	rmsd (Å)
derp5_dimer[3MQ1_A]	-1188	1.9
blot5_dimer <sub>6</sub>	-1152	>10
derf5_dimer <sub>5</sub>	-1353	3.0
suim5_dimer <sub>4</sub>	-1168	>10
lepd5_dimer <sub>3</sub>	-1190	>10
derp21_dimer <sub>6</sub>	-1266	>10
derf21_dimer <sub>2</sub>	-1252	>10
blot21_dimer <sub>7</sub>	-1574	8,86
derp5_V62AV65A	-1053	2.1
derp5_L69E	-1063	1.8
derp5_V62AV65AL69E	-930	2,1
derp5_L11AL93A	-1167.0	>10
derp5_L11AL93AV62AV65A	-1055.8	>10
derp5_all	-869	>10
derf5_V62AV65A	-1218	4.2
derf5_L69E	-1169	2.9
derf5_V62AV65AL69E	-1039	3.1
derf5_L11AL93A	-1250.6	4,71
derf5_L11L93V62AV65A	-1238.6	2,92
blot5_A62VA65V	-1266	>10
blot5_E69L	-1289	>10
blot5_A62VA65VE69L	-1401	>10
blot5_all	-1090	>10
blot5_dimer <sub>2</sub> _all	-1057	4.0
blot21_S62VY65V	-1552.7	8,67
blot21_S62VY65VE69L	-1494.7	>10
blot21_all	-1148,8	>10
blot21_I11LV62AV65A	-1444.7	>10

La structure prédite des dimères présentant les plus petits scores ClusPro de Sui m 5, Lep d 5, Blo t 5, Der p 21, Der f 21 et de Blo t 21 semble différente de celle de Der p 5 (3MQ1\_AB) (figure n°22); le meilleur dimère de Der p 21 ressemble beaucoup plus à celui de Der f 21 avec un rmsd de 4,01 Å. Le meilleur dimère de Sui m 5 est plus proche de celui de Lep d 5 avec un rmsd de 3,62 Å, et le meilleur dimère prédit pour Blo t 5 présente une structure totalement différente par rapport aux autres allergènes mais très similaire à celle de Blo t 21 avec un rmsd de 3,83 Å. En outre, parmi les solutions trouvées par ClusPro, nous avons trouvé des modèles de dimères de Sui m 5, Lep d 5, Der p 21 et Blo t 21 similaires à la structure expérimentale de Der p 5, mais leurs scores calculés par ClusPro sont plus élevés et donc moins favorables du point de vue énergétique (voir tableau A4, annexe II).



**Figure n°22:** Structures du dimère de Der p 5 (3MQ1\_AB) et des modèles de dimères prédites par ClusPro des allergènes des familles 5 et 21 étudiés. (a) Superposition de la structure dimérique de Der p 5 3MQ1\_AB (en rouge) et le modèle de dimère prédit de Der f 5 (en orange). (b) Superposition de structures entre les dimères prédit de Sui m 5 (vert) et de Lep d 5 (bleue). (c) Superposition de structures entre les dimères prédit de Der f 21 (jaune) et de Der p 21 (cyan). (d) Superposition de structures entre les dimères prédit de Blo t 5 (magenta) et de Blo t 21 (saumon).

Mueller et *al.* (2010a) ont discuté l'origine possible de la différence dans la propension de dimérisation entre Der p 5 et Blo t 5. Selon ces auteurs, la région de séquence GVRGV (numérotation de l'alignement 61-65) dans Der p 5 semble être importante dans la stabilité par la création des interactions de type valine zipper à l'interface des chaînes A et B de Der p 5. En effet, les deux valines V62 et V65 de la chaîne A sont impliqués dans des contacts hydrophobes avec les V65 et la V62 de la chaîne B, respectivement. Ces valines sont remplacés par des alanines dans Blo t 5, ce qui, par conséquent, devrait réduire la surface hydrophobe entre les deux chaînes et pourrait expliquer le fait que Blo t 5 reste sous forme monomérique. Ces résidus valines (V62 et V65) sont également remplacés par une sérine et une tyrosine, respectivement, dans Blo t 21.

Par ailleurs, le meilleur modèle de dimère de Der f 5 (*derf5\_dimer<sub>5</sub>*) est similaire à la structure du dimère de Der p 5. Ceci est en accord avec la grande identité de séquence entre les deux protéines, et avec la conservation des résidus impliqués dans les interactions à l'interface (tableaux n°10 et 12). En outre, la partie de séquence GVRGV (numérotation de l'alignement 61-65) dans Der p 5 est conservée uniquement dans Der f 5.

La différence de la nature des résidus en position 69 dans l'alignement pourrait également expliquer la différence entre Der p 5, Der f 5, Blo t 5 et Blo t 21 (figure n°19). Dans Der p 5 et Der f 5, cette position est occupée par une leucine impliquée dans des contacts hydrophobes. Par contre dans Blo t 5 et Blo t 21 on trouve un glutamate (chargé négativement).

Pour tester si les positions 62, 65 et 69 (selon la numérotation de l'alignement de séquences) jouent un rôle important dans la structure dimérique de Der p 5, nous avons muté les séquences des modèles de monomères de Blo t 5, Blo t 21, Der p 5 et Der f 5 et nous les avons soumis à ClusPro. Plus précisément, nous avons choisi les monomères qui ont conduit à la meilleure structure de dimère (*3MQ1\_A*, *derf5\_monomer<sub>5</sub>* [*3MQ1\_A*], *blot5\_monomer<sub>6</sub>*[*3MQ1\_A*] et *blot21\_monomer<sub>7</sub>*[*3MQ1\_A*]), introduit dans Der p 5 et dans Der f 5 les résidus qui se trouvent dans Blo t 5, et *vice versa*, et nous avons soumis ces structures mutées à ClusPro. Nous avons testé les mutations suivantes: V62A-V65A, L69E et V62A-V65A-L69E dans Der p 5 et dans Der f 5 (*derp5\_V62AV65A*, *derf5\_V62AV65A*, *derp5\_L69E*, *derf5\_L69E*, *derp5\_V62AV65AL69E*, *derf5\_V62AV65AL69E*, respectivement), et les mutations A62V-A65V, E69L et A62V-A65V-E69L dans Blo t 5

(blot5\_A62VA65V, blot5\_E69L, blot5\_A62VA65VE69L). Nous avons également introduit les mutations S62V-Y65V et S62V-Y65V-E69L dans Blo t 21 (blot21\_S62VY65V, blot21\_S62VY65VE69L) afin d'évaluer l'importance de ces résidus dans le mode de dimérisation de cet allergène.

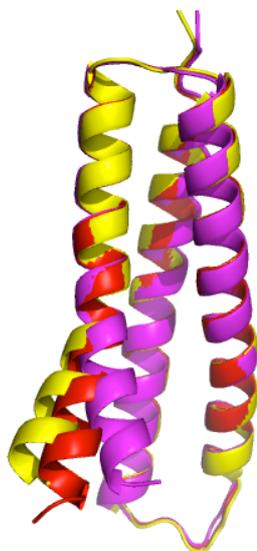
De plus, nous avons identifié toutes les positions qui sont impliquées dans les interactions à l'interface du dimère de Der p 5 (figure n°21) et nous avons introduit dans Blo t 5 et dans Blo t 21 les résidus qui se trouvent dans Der p 5. Les mutants ainsi créés ont été nommés "blot5\_all" et "blot21\_all", respectivement. Nous avons également introduit dans Der p 5 (au niveau de ces résidus d'interface) les résidus qui se trouvent dans Blo t 5; la structure ainsi mutée a été nommée "derp5\_all". Plus précisément, nous avons réalisé les mutations suivantes:

- dans blot5\_all: I12M-Q14R-A15I-N16H-A18Q-E20K-L27F-E35H-R50A-V54T-V55I-A62V-A65V-E69L-N76D-Y83L-E85M-L89S.
- dans blot21\_all: I11L-V12M-T14R-A15I-T16H-R18Q-F19I-H20K-I22G-L27F-H28Y-I29L-D35H-R47K-R50A-T53D-V54T-S55I-F58M-S62V-Y65V-F66L-E69L-F83L-A85M-T89S-L92I-L100E.
- dans derp5\_all: M12I-R14Q-I15A-H16N-Q18A-K20E-F27L-H35E-A50R-T54V-I55V-V62A-V65A-L69E-D76N-L83Y-M85E-S89L.

Les résultats trouvés sont présentés dans le tableau n°13. Les mutations V62A, V65A et L69E dans Der p 5 et dans Der f 5 ont conduit à la formation d'un dimère qui est toujours proche de la structure expérimentale de Der p 5: le rmsd entre le dimère prédit et la structure 3MQ1\_AB est inférieur ou égal à 3.1Å. Toutefois, l'augmentation des scores ClusPro suggère que ces dimères sont moins stables. Donc, ces positions 62, 65 et 69 semblent jouer un rôle dans la stabilité du dimère, mais, selon nos simulations, leur mutation ne serait pas suffisante pour empêcher la formation du dimère. La mutation des résidus d'interface de Der p 5 avec leurs correspondant dans Blo t 5 (derp5\_all) n'a pas donné une structure dimérique de Der p 5 similaire à sa structure native (3MQ1\_AB), avec une valeur de score ClusPro plus élevée.

Les mutations effectuées au niveau de la séquence de Blo t 5 (A62V, A65V et E69L) et de Blo t 21 (S62V, Y65V et E69L) ne favorisent pas la formation d'une structure dimérique similaire à celle de Der p 5, le rmsd calculé entre ces structures mutées et le dimère de Der p 5 (3MQ1\_AB) reste supérieur à 10 Å (tableau n°13). Blot5\_all et blot21\_all contiennent tous

les résidus qui sont impliqués dans les interactions à l'interface de Der p 5. Nous étions surpris de constater que la structure des dimères obtenus avec blot5\_all et blot21\_all soient différentes de celle de Der p 5 (rmsd très grand) (tableau n°13). L'analyse des structures des modèles de monomères de Blo t 5 a révélé que l'hélice  $N_{terminale}$  est moins coudée dans la structure nommée blot5\_monomer<sub>6</sub>[3MQ1\_A] (figure n°23), ce qui entrave la formation d'un dimère enchevêtré comme celui de Der p 5. L'hélice est plus coudée dans la structure blot5\_monomer<sub>2</sub>[3MQ1\_A] (figure n°23, et tableau A3 en annexe II).



**Figure n°23:** Représentation de l'hélice  $N_{terminale}$  coudée après superposition de la structure expérimentale de Der p 5 (3MQ1\_A) en rouge avec les structures prédites de Blo t 5: blot5\_monomer<sub>6</sub>[3MQ1\_A] en magenta et blot5\_monomer<sub>2</sub>[3MQ1\_A] en jaune.

Nous avons donc introduit toutes les mutations mentionnées ci-dessus (I12M-Q14R-A15I-N16H-H17E-A18Q-E20K-H24L-E35H-R50A-V54T-V55I-A62V-A65V-E69L-N76D-Y83L-E85M-L89S) dans cette structure (blot5\_monomer<sub>2</sub>[3MQ1\_A]), et nous l'avons soumise à ClusPro. La meilleure structure de dimère (blot5\_dimer<sub>2</sub>\_all) trouvée présente un rmsd de 4,0 Å avec la structure expérimentale Der p 5 (tableau n°13). Il est donc nécessaire de muter un grand nombre de positions de séquence pour induire chez Blo t 5 la formation d'une structure de dimère similaire à celle de Der p 5.

D'autre part, l'analyse de la structure de Der p 5 nous a montré la présence de contacts hydrophobes entre les leucines en positions 11 et 93 de notre alignement de séquence. Ces derniers pourraient jouer un rôle également dans la stabilité du dimère de Der p 5 en formant

des interactions de type leucine-zipper. Des études préalables sur d'autres protéines ont montré que la mutation de ce genre de contact qui se trouvent à l'extrémité *C\_terminale* de la protéine conduit à l'abolissement de l'homodimérisation de cette dernière (Kim et *al.*, 2001). Ce qui est intéressant est que la mutation de ces résidus leucines dans la structure 3MQ1\_A de Der p 5 a conduit à l'obtention par ClusPro d'un homodimère qui est différent de 3MQ1\_AB (tableau n°13). La mutation des résidus de contacts hydrophobes L11, L93, V62 et V65 de Der p 5 en alanine a empêché la formation de la structure enchevêtrée de cet allergène. Ce résultat montre réellement l'importance de ces résidus d'interface dans Der p 5. Cependant, même si ces résidus sont tous conservé dans Der f 5, leur mutation n'était pas suffisante pour inhiber la formation d'un dimère similaire à celui de 3MQ1\_AB de Der p 5.

*Chapitre III:*

*Conclusions*

### III. 4 Conclusions

Le processus de dimérisation est l'une des caractéristiques majeures de l'activité biologique des allergènes (Schöll et *al.*, 2005), soit par sa contribution à la réticulation des IgE sur les mastocytes et les basophiles ou par la création des cavités qui permettent la fixation des ligands et l'activation du système immunitaire.

Parmi les allergènes des familles 5 et 21, seules les structures de Blo t 5, Blo t 21 et Der p 5 ont été déterminées expérimentalement (Naik et *al.*, 2008, Tan et *al.*, 2012; Mueller et *al.*, 2010a). Les structures de Blo t 5 et de Blo t 21 (codes PDB 2JMH et 2LM9, respectivement) sont sous formes de monomères tandis que celle de Der p 5 a été trouvée comme étant un homodimère (deux monomères enchevêtrés grâce à leurs hélices  $N_{terminale}$  coudées), formant une cavité hydrophobe au centre de la molécule (chaînes A et B du fichier PDB 3MQ1).

Dans cette étude, nous avons utilisé des approches de modélisation moléculaire pour analyser la possibilité que d'autres allergènes d'acariens des familles 5 et 21 (Der f 5, Sui 5 m, Lep d 5, Der p 21 et Der f 21) puissent adopter une structure de dimère similaire à Der p 5, avec une grande cavité hydrophobe.

La modélisation de la structure monomérique des allergènes étudiés avec comme patrons la structure de la chaîne A de Der p 5 (hélice  $N_{terminale}$  coudée), de Blo t 21 (hélice  $N_{terminale}$  légèrement coudée) et de Blo t 5 (hélice  $N_{terminale}$  droite) suggère que Der f 5 a une préférence pour la structure avec une hélice coudée, tandis que la conclusion est moins claire pour les autres allergènes.

Les monomères modélisés avec une hélice  $N_{terminale}$  coudée ont été soumis au programme ClusPro en vue de prédire la structure du dimère de ces allergènes. Ce type de programme est incapable de prédire si une protéine formera ou pas un dimère *in vitro*, mais il nous a permis d'évaluer si la structure de l'éventuel dimère prédit avec le meilleur score ClusPro pourrait être similaire à celle de Der p 5. Les résultats trouvés montrent que seul Der f 5 peut former un dimère comme celui de Der p 5, avec une cavité hydrophobe d'environ 2000 Å<sup>3</sup>.

Mueller et *al.* (2010a) ont suggéré que les différences entre Blo t 5 et Der p 5 en positions de séquences 62 et 65 (selon la numérotation de l'alignement de séquences) pourrait expliquer le fait que Blo t 5 soit un monomère et Der p 5 un dimère. L'alignement de séquences révèle que

la position 69 pourrait également participer à la différence entre les séquences de ces deux protéines: une leucine est présente dans Der p 5 et un glutamate est trouvé dans Blo t 5 et dans Blo t 21. En outre, ces trois positions (V62, V65 et L69) de séquence sont conservées uniquement entre Der p 5 et Der f 5. La mutation de ces trois résidus dans Der p 5 avec leurs correspondants dans Blo t 5 suggère que ces résidus pourraient jouer un rôle dans la stabilité du dimère. Mais ces mutations ne sont pas suffisantes pour éviter la formation de la structure du dimère de Der p 5 et promouvoir sa formation pour Blo t 5 ou pour Blo t 21. Cependant, d'après l'analyse de la structure de Der p 5, nous pouvons suggérer l'importance des résidus leucines en positions 11 et 93 (de notre alignement de séquence) dans la stabilité et la formation du dimère enchevêtré de cet allergène. La mutation de ces deux résidus a permis l'abolissement de la formation de la structure native de Der p 5. Ceci montre l'importance de ces interactions hydrophobes de type leucine-zipper, à côté des valines-zipper, dans la structure de dimère de Der p 5.

Nous avons également identifié les positions de séquences qui sont impliquées dans les interactions à l'interface du dimère de Der p 5 et nous avons analysé leur conservation dans Blo t 5, Blo t 21 et dans Der f 5. Ce qui est intéressant, est que la mutation de Blo t 5 en les résidus d'interface de Der p 5 peut conduire à une structure de dimère enchevêtrée avec une cavité de 1585 Å<sup>3</sup>. Ceci dit, ce résultat suggère qu'il est nécessaire de muter un grand nombre de résidus pour favoriser la formation d'une structure enchevêtrée du dimère de Blo t 5.

*Conclusion générale et  
perspectives*

Ce sujet de thèse se place dans la thématique générale de l'allergie aux acariens, en particulier des allergies dues aux espèces *Dermatophagoides pteronyssinus*, *Dermatophagoides farinae* et *Blomia tropicalis*. A l'heure actuelle, plusieurs allergènes d'acariens ont été isolés. Ils sont caractérisés par (1) leurs épitopes (déterminants antigéniques) qui représentent leur site de reconnaissance et de liaison avec les anticorps spécifiques de type IgE et (2) par leur fonction qui contribue activement à l'aggravation des symptômes de l'allergie.

L'objectif de cette thèse a consisté en l'utilisation d'approches bioinformatiques afin d'étudier la structure et la fonction des allergènes des familles 5 et 21, en particulier les allergènes Der p 5 et Der p 21 de *Dermatophagoides pteronyssinus* Der f 5 et Der f 21 de *Dermatophagoides farinae*, Blo t 5 et Blo t 21 de *Blomia tropicalis*, Sui m 5 de *Suidasia medanensis* et Lep d 5 de *Lepidoglyphus destructor*. Le choix de ces allergènes résulte des faits suivants: (i) ils ont été identifiés comme jouant un rôle important et déterminant dans l'allergie aux acariens, (ii) très peu de choses sont connues à propos de leurs structures 3D, leur réactivité croisée et leur fonction biologique, et (iii) la petite taille de ces protéines en fait des bons candidats à la modélisation *in silico*.

Parmi ces allergènes des familles 5 et 21, seules les structures de Blo t 5 (code PDB 2JMH, Naik et *al.*, 2008), de Blo t 21 (code PDB 2LM9, Tan et *al.*, 2012) et de Der p 5 (code PDB 3MQ1, Mueller et *al.*, 2010) ont été résolues soit par RMN ou par cristallographie aux rayons-X. En outre, peu d'études ont été faites sur la détermination de leurs épitopes et de leur réactivité croisée. Au sein de ces deux familles, seulement deux sites antigéniques dans Blo t 5 (Naik et *al.*, 2008; Chan et *al.*, 2008) et un seul dans Blo t 21 sont connus. Les épitopes sont des régions peptidiques en surface des antigènes qui se fixent spécifiquement aux anticorps de type IgE et sont donc responsables de l'immunoréactivité et de l'activation de la réponse immunitaire humorale contre ces allergènes. La conservation de ces régions antigéniques dans différents allergènes est la raison pour laquelle ces derniers pourraient présenter une réactivité croisée entre eux. La recherche de ces sites antigéniques est donc primordiale afin de comprendre l'allergénicité de ces antigènes et permettre la conception rationnelle des variants hypoallergéniques utilisés comme des produits de désensibilisation en immunothérapie spécifique.

Certains allergènes présentent des activités et des fonctions biologiques qui pourraient aggraver la réponse immunitaire et les symptômes de l'allergie. C'est le cas de l'activité

protéase des allergènes majeurs d'acariens des familles 1, 3 et 9 et le transport des ligands hydrophobes par les allergènes de famille 2 (Der p 2 et Der f 2). Une autre particularité de certains allergènes est leur phénomène de dimérisation qui pourrait contribuer à la réticulation des IgE (Schöll et *al.*, 2005) ou créer des cavités et permettre ainsi le transport des ligands. La compréhension de la fonction biologique et les propriétés d'un allergène passe donc par l'étude de sa structure tridimensionnelle (3D).

Deux objectifs ont été tracés dans notre étude. Dans le premier objectif, nous avons cherché à caractériser les épitopes de Der p 5. Nous avons plus particulièrement analysé les épitopes déterminés expérimentalement de Blo t 5 et de Blo t 21 avec leur correspondants dans les allergènes étudiés et nous avons pu suggérer que la faible réactivité croisée trouvées entre ces allergènes de familles 5 et 21 pourrait être due en partie à la conservation de la région épitope située dans la deuxième boucle et l'hélice  $C_{terminale}$  (appelé épitope-I dans notre deuxième chapitre).

La majorité des épitopes sont conformationnels (Pomés, 2010; El-Manzalawy et Honavar, 2010), donc dépendants de la structure 3D des allergènes. La résolution de la structure 3D de Der p 5 facilite donc la discussion et la recherche de ses sites antigéniques, par la comparaison des résidus épitopes déterminés expérimentalement dans Blo t 5 et dans Blo t 21, et surtout par la prédiction de ses épitopes par l'utilisation d'un programme qui prend en compte les conformations et les propriétés structurales des épitopes.

Sur la base de ses performances, le programme "SEPPA" a été utilisé dans notre travail pour la prédiction des épitopes B de Der p 5. Dans un premier temps, SEPPA nous a permis de définir quatre régions antigéniques pour Der p 5: L11-H16 (épitope-Nt), H35-D46 (épitope-I), L69-E84 (épitope-II) et K106-V110 (épitope-Ct) (selon la numérotation dans notre alignement de séquences multiple entre les allergènes des familles 5 et 21). Par la suite, des résidus appartenant à ces régions et qui pourraient jouer un rôle crucial dans la fixation aux IgE ont été choisis selon nos critères de sélection (nature des acides aminés et leur accessibilité au solvant) pour la validation expérimentale dans le laboratoire du Centre d'Ingénierie des Protéines à Liège (Belgique). En outre, l'impact sur la stabilité thermodynamique de la mutation de ces résidus en alanine a été évalué par le programme PoPMuSiC. Les résultats suggèrent que l'alanine ne déstabilise pas de manière importante la structure de Der p 5 lorsqu'elle est introduite à ces positions. Les protéines mutées ont été

exprimées par nos collaborateurs et leur réponse allergique sera ensuite évaluée expérimentalement (cette étape est encore en cours de validation).

L'importance de la prédiction des épitopes B de Der p 5 s'inscrit dans la voie de recherche des hypoallergènes. L'objectif est de réaliser la synthèse de nouvelles molécules (appelées hypoallergènes) qui conservent la possibilité d'agir sur les lymphocytes T mais qui ne peuvent pas se lier aux IgE. Ce travail contribue donc à l'étude du mécanisme de l'allergénicité de l'allergène Der p 5 et pourrait aboutir à la conception de ses variants hypoallergéniques pour l'immunothérapie.

Le deuxième objectif de la thèse s'est porté sur l'étude de la structure 3D des allergènes des familles 5 et 21 (Der p 5, Der f 5, Sui m 5, Lep d 5, Der p 21, Der f 21, Blo t 5 et Blo t 21). Seules les structures de Blo t 5, Blo t 21 et de Der p 5 ont été déterminées expérimentalement. Cependant, les structures 3D de Blo t 5 et de Blo t 21 sont sous forme de monomères tandis que celle de Der p 5 a été trouvée comme étant un homodimère formant une cavité hydrophobe au centre de la molécule. La comparaison entre ces structures montre une différence au niveau de la flexibilité de leur hélice  $N_{terminale}$  qui est coudée dans Der p 5 lui permettant ainsi la formation de son dimère avec les deux chaînes enchevêtrées et la création d'une grande cavité hydrophobe au centre de la molécule. Notre travail a été tracé dans le but d'étudier *in silico* la propension de dimérisation de ces allergènes des familles 5 et 21 (Der f 5, Sui m 5, Lep d 5, Der p 21, Der f 21, Blo t 5 et Blo t 21) à adopter des dimères semblables à celui de Der p 5.

D'après nos résultats, seuls Der p 5 et Der f 5 pourraient présenter une structure de dimère similaire, avec la formation d'une grande cavité pouvant fixer des ligands hydrophobes. Selon nos simulations, les autres allergènes étudiés Blo t 5, Sui 5 m, Lep d 5, Der p 21 et Der f 21 ne présentent pas une telle propension.

Par ailleurs, nous avons également identifié des mutations qui pourraient déstabiliser ou entraver la dimérisation de Der p 5 et probablement de Der f 5. En effet, l'oligomérisation de Der p 5 est dépendante de la concentration (Mueller et *al.*, 2010a), donc des mutations qui diminuent la stabilité du dimère devraient augmenter la concentration nécessaire pour observer ces dimères. Par conséquent, la caractérisation des mutations proposées par notre étude (au niveau des résidus de l'interface du dimère de Der p 5) pourrait aider à comprendre

si la formation d'une cavité hydrophobe par la dimérisation de Der p 5 joue un rôle important dans l'induction d'une réaction allergique, et si Der f 5 présente cette particularité.

En outre, il serait surprenant que le mécanisme de l'induction de la réponse immunitaire contre ces allergènes des familles 5 et 21 soit différent. D'après nos simulations, seul Der f 5 pourrait former un dimère similaire à celui de Der p 5. Par conséquent, nous ne pensons pas que le processus de dimérisation de Der p 5 et peut-être de Der f 5 soit impliqué par lui-même dans l'induction de la réponse allergique. Toutefois, l'ouverture de la structure par la formation d'un coude dans l'hélice N<sub>terminale</sub>, même si ce changement de conformation est transitoire, pourrait provoquer l'exposition de certains acides aminés hydrophobes et donc la liaison des ligands hydrophobes et induire la réponse allergique. La particularité de Der p 5 et probablement de Der f 5 serait que la structure du monomère avec l'hélice N<sub>terminale</sub> coudée pourrait être stabilisée par sa dimérisation.

Par ailleurs, l'existence de structures communes entre ces allergènes pourrait non seulement contribuer à la compréhension de leurs propriétés mais surtout permettre de caractériser des épitopes communs et montrer l'implication de la réactivité croisée, notamment entre Der p 5 et Der f 5 dans une part importante de l'allergénicité des acariens de poussières de maisons.

La particularité de fixation des ligands dans des cavités hydrophobes est partagée avec d'autres allergènes d'acariens. Les exemples les plus étudiés sont les allergènes d'acarien de la famille 2 (Der p 2 et Der f 2). Ces derniers présentent une similarité structurale avec la protéine MD2, qui joue un rôle dans la reconnaissance des LPS et dans leur présentation aux récepteurs TLR4, ce qui déclenche la réponse immunitaire innée. Par ailleurs, la résonance magnétique nucléaire et les dosages biochimiques ont indiqué que Der p 7 (l'allergène de famille 7 de *Dermatophagoides pteronyssinus*) présente une cavité hydrophobe, ne lie pas les LPS, mais lie avec une faible affinité le lipopeptide polymyxine B d'origine bactérienne (*Bacillus polymixa*) (Mueller et al., 2010b). De plus, Valerio et al. (2005) ont séquencé, à partir des extraits d'acariens, des gènes d'ARN 16S de certaines bactéries Gram négatives (certaines espèces de *Bartonella*, *Escherichia coli*, certaines espèces de *Pseudomonas* et de *Acinetobacter*) qui se trouvent être la source des endotoxines trouvées en complexe avec les allergènes des acariens et qui aggravent l'allergie.

Tous ces résultats suscitent de nouvelles interrogations, et ouvrent donc la voie à des études futures dans le volet bioinformatique et/ou expérimental.

Une perspective intéressante et prometteuse sera le criblage des ligands hydrophobes et l'étude de leur liaison dans la cavité de Der p 5 en utilisant des programmes d'amarrage protéine-ligands. Cette étude permettra la compréhension de la fonction biologique de ces allergènes et le mécanisme par lequel ils pourraient déclencher et/ou exacerber la réponse immunitaire. Nous pourrions également envisager de faire une étude de la dynamique moléculaire de la flexibilité des l'hélice N<sub>terminale</sub> de Der p 5 et de Der f 5 qui pourrait apporter plus de lumière quant au mécanisme de dimérisation de Der p 5 et probablement de Der f 5. Les hypothèses émises par ces approches *in silico* pourront ensuite être testées expérimentalement.

Enfin, il serait également nécessaire de caractériser la structure et les épitopes de Der f 5. Les résultats de ces études pourraient éclairer nos connaissances quant à la réactivité croisée entre ces deux allergènes et aider à la conception rationnelle des vaccins hypoallergéniques pour l'immunothérapie spécifique contre ces allergènes.

Par ailleurs, tous les résultats obtenus dans cette thèse ont fait l'objet de plusieurs communications nationales et internationales (liste en annexe III), d'un article publié dans le *Journal of Biomolecular Structure and Dynamics* (avec un facteur d'impact de 4,9) (annexe III) et un prix m'a été décerné lors du 5ème Congrès de Médecine et de Chirurgie à Tlemcen (**Algérie**).

*Références  
bibliographiques*

**A**

- Abed-Benamara, M., Fain, A., & Abed, L. (1983). Note préliminaire sur la faune acarologique des poussières de matelas d'Alger. *Acarologia*, 79–83. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Note+préliminaire+sur+la+faune+acarologique+des+poussières+de+matelas+d'Alger#0>
- Adam, E., Hansen, K.K, Astudillo Fernandez, O., Coulon, L., Bex, F., Duhant, X., Jaumotte, E., Hollenberg, M. D., & Jacquet, A. (2006). The house dust mite allergen Der p 1, unlike Der p 3, stimulates the expression of interleukin-8 in human airway epithelial cells via a proteinase-activated receptor-2-independent mechanism. *The Journal of biological chemistry*, 281(11), 6910–23. doi:10.1074/jbc.M507140200
- Aki, T., Ono, K., Paik, S. Y., Wada, T., Jyo, T., Shigeta, S., Murooka, Y., & Oka, S. (1994). Cloning and characterization of cDNA coding for a new allergen from the house dust mite, *Dermatophagoides farinae*. *International archives of allergy and immunology*, 103(4), 349–56. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7510558>
- Aki, T., Kodama, T., Fujikawa, A., Miura, K., Shigeta, S., Wada, T., Jyo, T., Murooka, Y., Oka, S., & Ono, K. (1995). Immunochemical characterization of recombinant and native tropomyosins as a new allergen from the house dust mite, *Dermatophagoides farinae*. *Journal of Allergy and Clinical Immunology*, 96(1), 74–83. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0091674995700358>
- Alix, A. J. (1999). Predictive estimation of protein linear epitopes by using the program PEOPLE. *Vaccine*, 18(3-4), 311–4. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10506656>
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17), 3389–402. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=146917&tool=pmcentrez&rendertype=abstract>
- Andersen, P. H., Nielsen, M., & Lund, O. (2009). Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Science*, 15, 2558–2567. doi:10.1110/ps.062405906.2558
- Ansari, H. R., & Raghava, G. P. (2010). Identification of conformational B-cell Epitopes in an antigen from its primary sequence. *Immunome research*, 6, 6. doi:10.1186/1745-7580-6-6
- Arbes, S. J., Cohn, R. D., Yin, M., Muilenberg, M. L., Burge, H. a., Friedman, W., & Zeldin, D. C. (2003). House dust mite allergen in US beds: Results from the first national

- survey of lead and allergens in housing. *Journal of Allergy and Clinical Immunology*, 111(2), 408–414. doi:10.1067/mai.2003.16
- Arlian, L. G., & Platts-Mills, T. a. E. (2001). The biology of dust mites and the remediation of mite allergens in allergic disease. *Journal of Allergy and Clinical Immunology*, 107(3), S406–S413. doi:10.1067/mai.2001.113670
- Arlian, L. G. (2002). Arthropod allergens and human health. *Annual Review of Entomology*, 47, 395–433. Retrieved from [http://www.annualreviews.org/doi/full/10.1146/annurev.ento.47.091201.145224?url\\_ver=Z39.88-2003&rft\\_id=ori:rid:crossref.org&rft\\_dat=cr\\_pub=pubmed](http://www.annualreviews.org/doi/full/10.1146/annurev.ento.47.091201.145224?url_ver=Z39.88-2003&rft_id=ori:rid:crossref.org&rft_dat=cr_pub=pubmed)
- Arlian, L. G., Morgan, M. S., Vyszynski-Moher, D. L., & Sharra, D. (2009). Cross-reactivity between storage and dust mites and between mites and shrimp. *Experimental & applied acarology*, 47(2), 159–72. doi:10.1007/s10493-008-9199-x
- Arruda, L. K., Vailes, L. D., Platts-Mills, T. A., Fernandez-Caldas, E., Montealegre, F., Lin, K. L., Chua, K. Y., Rizzo, M. C., Naspitz, C. K., & Chapman, M. D. (1997). Sensitization to *Blomia tropicalis* in patients with asthma and identification of allergen Blo t 5. *American journal of respiratory and critical care medicine*, 155(1), 343–50. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9001334>
- Asturias, J. A., Arilla, M. C., Gómez-Bayón, N., Martínez, A., Martínez, J., & Palacios, R. (1998). Sequencing and high level expression in *Escherichia coli* of the tropomyosin allergen (Der p 10) from *Dermatophagoides pteronyssinus*. *Biochimica et biophysica acta*, 1397(1), 27–30. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9545522>
- Asturias, J. A., Ibarrola, I., Arilla, M. C., Vidal, C., Ferrer, A., Gamboa, P.M., Viñuela, J. E., Sanz, M.L., Andreu, C., & Martínez, A. (2009). Engineering of major house dust mite allergens Der p 1 and Der p 2 for allergen-specific immunotherapy. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*, 39(7), 1088–98. doi:10.1111/j.1365-2222.2009.03264.x

## B

- Baker, D., & Sali, a. (2001). Protein structure prediction and structural genomics. *Science (New York, N.Y.)*, 294(5540), 93–6. doi:10.1126/science.1065659
- Baqueiro, T., Carvalho, F. M., Rios, C. F., dos Santos, N. M., & Alcântara-Neves, N. M. (2006). Dust mite species and allergen concentrations in beds of individuals belonging to different urban socioeconomic groups in Brazil. *The Journal of asthma: official journal of the Association for the Care of Asthma*, 43(2), 101–5. doi:10.1080/02770900500497958

- Barlow, D., Edwards, M., & Thornton, J. (1986). Continuous and discontinuous protein antigenic determinants. *Nature*, 322, 747–748. Retrieved from <http://www.nature.com/nature/journal/v322/n6081/abs/322747a0.html>
- Barre, A., Borges, J.-P., Culerrier, R., & Rougé, P. (2005). Homology modelling of the major peanut allergen Ara h 2 and surface mapping of IgE-binding epitopes. *Immunology letters*, 100(2), 153–8. doi:10.1016/j.imlet.2005.03.014
- Barrio, M. A., Soeria-Atmadja, D., Nistér, A., Gustafsson, M. G., Hammerling, U., & Bongcam-Rudloff, E. (2007). EVALLER: a web server for in silico assessment of potential protein allergenicity. *Nucleic acids research*, 35(Web Server issue), W694–700. doi:10.1093/nar/gkm370
- Bedada, G. B., Heinrich, J., Götschi, T., Downs, S. H., Forsberg, B., Jarvis, D., Luczynska, C., Soon, A., Sunyer, J., Toren, K., & Künzli, N. (2007). Urban background particulate matter and allergic sensitization in adults of ECRHS II. *International journal of hygiene and environmental health*, 210(6), 691–700. doi:10.1016/j.ijheh.2006.11.005
- Belhocine, M., & Ait-Khaled, N. (1991). Prévalence de l'asthme dans une région d'Algérie. *Bulletin de l'Union internationale contre la tuberculose et les maladies respiratoires*, 66, 97–100. Retrieved from <http://cat.inist.fr/?aModele=afficheN&cpsid=4323586>
- Bennett, B., & Thomas, W. (1996). Cloning and sequencing of the group 6 allergen of *Dermatophagoides pteronyssinus*. *Clinical & Experimental Allergy*, 10, 1150–4. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2222.1996.tb00501.x/abstract>
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2011). GenBank. *Nucleic acids research*, 39(Database issue), D32–7. doi:10.1093/nar/gkq1079
- Berman, H. M., Bhat, T. N., Bourne, P. E., Feng, Z., Gilliland, G., Weissig, H., & Westbrook, J. (2000). The Protein Data Bank and the challenge of structural genomics. *Nature structural biology*, 7 Suppl(November 2000), 957–9. doi:10.1038/80734
- Bidad, K., Nicknam, M. H., & Farid, R. (2011). A review of allergy and allergen specific immunotherapy. *Iranian journal of allergy, asthma, and immunology*, 10(1), 1–9. doi:10.1016/j.ijaai.19
- Björkstén, B., Clayton, T., Ellwood, P., Stewart, A., & Strachan, D. (2008). Worldwide time trends for symptoms of rhinitis and conjunctivitis: Phase III of the International Study of Asthma and Allergies in Childhood. *Pediatric allergy and immunology: official publication of the European Society of Pediatric Allergy and Immunology*, 19(2), 110–24. doi:10.1111/j.1399-3038.2007.00601.x

- Blythe, M., & Flower, D. (2005). Benchmarking B cell epitope prediction: underperformance of existing methods. *Protein Science*, *14*, 246–248. doi:10.1110/ps.041059505.Four
- Bouayad, Z., & Afif, H. (1998). L'épidémiologie de l'asthme et de la rhinite dans les pays au Sud de la Méditerranée. *Revue Française d'Allergologie et d'Immunologie Clinique*, *38*. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0335745798800919>
- Boumendjel, A., Tridon, A., & Ughetto, S. (2010). Environnement allergénique d'une population d'enfants asthmatiques à Annaba (Algérie). *Annales de Biologies Cliniques*, *68*(3), 317–24. Retrieved from <http://www.jle.com/e-docs/00/04/57/60/>
- Bousquet, J., Jeffery, P.K., Busse, W.W., Johnson, M., & Vignola, A. M. (2000). Asthma From Bronchoconstriction to Airways Inflammation and Remodeling. *American journal of respiratory and critical care medicine*, *161*, 1720–1745. Retrieved from <http://ajrccm.atsjournals.org/content/161/5/1720.short>
- Bousquet, P.-J., Chinn, S., Janson, C., Kogevinas, M., Burney, P., & Jarvis, D. (2007). Geographical variation in the prevalence of positive skin tests to environmental aeroallergens in the European Community Respiratory Health Survey I. *Allergy*, *62*(3), 301–9. doi:10.1111/j.1398-9995.2006.01293.x
- Bousquet, P.-J., Leynaert, B., Neukirch, F., Sunyer, J., Janson, C. M., Anto, J., Jarvis, D., & Burney, P. (2008a). Geographical distribution of atopic rhinitis in the European Community Respiratory Health Survey I. *Allergy*, *63*(10), 1301–9. doi:10.1111/j.1398-9995.2008.01824.x
- Bousquet, J., Khaltaev, N., & Cruz, A. (2008b). Allergic Rhinitis and its Impact on Asthma (ARIA) 2008. *Allergy*, *63*, 8–160. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1398-9995.2007.01620.x/full>
- Bousquet, J., Schünemann, H.J., Samolinski, B., Demoly, P., Baena-Cagnani, C.E., Bachert, C., Bonini, S., Boulet, L.P., Bousquet, P.J., Brozek, J.L., Canonica, G.W., Casale, T.B., Cruz, A.A., Fokkens, W.J., Fonseca, J.A., van, T., & { W. H. O. C. C. for A. and R. }. (2012). Allergic Rhinitis and its Impact on Asthma (ARIA) achievements in 10 years and future needs. *Journal of Allergy and Clinical Immunology*, *130*(5), 1049–62. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/23141389>
- Bouziane, F., Benaissa, A., & Abed-Benamara, M. (2001). VIème JOURNEE NATIONALE DE PARASITOLOGIE. *Algérie*.
- Brusic, V., Millot, M., & Petrovsky, N. (2003). Allergen databases. *Allergy*, 1093–1100. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1034/j.1398-9995.2003.00248.x/full>

Bui, H-H., Sidney, J., Li, W., Fusseder, N., & Sette, A. (2007). Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC bioinformatics*, 8, 361. doi:10.1186/1471-2105-8-361

## C

Camacho, C., Ma, H., & Champ, P. (2006). Scoring a diverse set of high-quality docked conformations: A metacore based on electrostatic and desolvation interactions. *Proteins*, 11–13. doi:10.1002/prot.v63

Campbell, S. J., Gold, N. D., Jackson, R. M., & Westhead, D. R. (2003). Ligand binding: functional site location, similarity and docking. *Current Opinion in Structural Biology*, 13(3), 389–395. doi:10.1016/S0959-440X(03)00075-7

Caraballo, L., Avjioglu, A., Marrugo, J., Puerta, L., & Marsh, D. (1996). Cloning and expression of complementary DNA coding for an allergen with common antibody-binding specificities with three allergens of the house dust mite *Blomia tropicalis*. *Journal of Allergy and Clinical Immunology Allergy Clin Immunol*, 98, 573–579.

Caraballo, L., Mercado, D., Jiménez, S., Moreno, L., Puerta, L., & Chua, K. Y. (1998). Analysis of the cross-reactivity between BtM and Der p 5, two group 5 recombinant allergens from *Blomia tropicalis* and *Dermatophagoides pteronyssinus*. *International archives of allergy and immunology*, 117(1), 38–45. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9751846>

Caruso, G., Damiani, V., Salerni, L., & Passali, F. M. (2009). Atopy: pediatric ENT manifestations in children. *International journal of pediatric otorhinolaryngology*, 73 Suppl 1, S19–25. doi:10.1016/S0165-5876(09)70005-9

Chambers, J., Thind, B., Dunn, J., & Pearson, D. (1999). The importance of storage mite allergens in occupational and domestic environments. *Proceedings of the 3rd International Conference on Urban Pests*, 559–569. Retrieved from <http://www.icup.org.uk/reports\ICUP472.pdf>

Chan, S. L., Ong, S. T., Ong, S. Y., Chew, F. T., & Mok, Y. K. (2006). Nuclear magnetic resonance structure-based epitope mapping and modulation of dust mite group 13 allergen as a hypoallergen. *Journal of immunology (Baltimore, Md. : 1950)*, 176(8), 4852–60. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16585580>

Chan, S. L., Ong, T. C., Gao, Y. F., Tiong, Y. S., Wang, D. Y., Chew, F. T., & Mok, Y. K. (2008). Nuclear magnetic resonance structure and IgE epitopes of Blo t 5, a major dust mite allergen. *Journal of immunology (Baltimore, Md. : 1950)*, 181(4), 2586–96. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/18684949>

- Chapman, M., Heymann, P., & Platts-Mills, T. (1987). Epitope mapping of two major inhalant allergens, Der p I and Der f I, from mites of the genus *Dermatophagoides*. *Journal of immunology (Baltimore, Md. : 1950)*, 139(5), 1479–84. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2442247>
- Chapman, M., Pomés, A., Breiteneder, H., & Ferreira, F. (2007). Nomenclature and structural biology of allergens. *The Journal of allergy and clinical immunology*, 119(2), 414–20. doi:10.1016/j.jaci.2006.11.001
- Chapman, M. (2008). Allergen nomenclature. *Clinical allergy and immunology*, 21(July), 47–58. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/18828497>
- Chapman, M., Pomés, A., & Aalberse, R. C. (2009). Molecular Biology of Allergens: Structure and Immune Recognition. (R. Pawankar, S. T. Holgate, & L. J. Rosenwasser, Eds.) *Allergy Frontiers: Epigenetics, Allergens and Risk Factors*, 1, 265–289. doi:10.1007/978-4-431-72802-3
- Chen, R., Li, L., & Weng, Z. (2003). ZDOCK: an initial-stage protein-docking algorithm. *Proteins*, 52(1), 80–7. doi:10.1002/prot.10389
- Chen, Z., Li, M., Chen, Y., Ji, J., & Li, Y. (2009). Effects of *Dermatophagoides pteronyssinus* allergen-specific immunotherapy on the serum interleukin-13 and pulmonary functions in asthmatic children. *Chin Med J*, 122(2008), 1157–1161. doi:10.3760/cma.j.issn.0366-6999.2009.10.008
- Chen, W. H., Sun, P. P., Lu, Y., Guo, W. W., Huang, Y. X., & Ma, Z. Q. (2011). MimoPro: a more efficient Web-based tool for epitope prediction using phage display libraries. *BMC bioinformatics*, 12, 199. doi:10.1186/1471-2105-12-199
- Chen, K.-W., Blatt, K., Thomas, W.R., Swoboda, I., Valent, P., Valenta, R., & Vrtala, S. (2012a). Hypoallergenic Der p 1/Der p 2 combination vaccines for immunotherapy of house dust mite allergy. *The Journal of allergy and clinical immunology*, 130(2), 435–43.e4. doi:10.1016/j.jaci.2012.05.035
- Chen, K.-W., Focke-Tejkl, M., Blatt, K., Kneidinger, M., Gieras, A., Dall'Antonia, F., Faé, I., Fischer, G., Keller, W., Valent, P., Valenta, R., & Vrtala, S. (2012b). Carrier-bound nonallergenic Der p 2 peptides induce IgG antibodies blocking allergen-induced basophil activation in allergic patients. *Allergy*, 67(5), 609–21. doi:10.1111/j.1398-9995.2012.02794.x
- Cheong, N., Soon, S., & Ramos, J. (2003). Lack of human IgE cross-reactivity between mite allergens Blo t 1 and Der p 1. *Allergy*, 912–920. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1034/j.1398-9995.2003.00215.x/full>

- Cheong, Nge, Ramos, J. D. A, Tang, C. Y., Chng, H. H., Yao, R., Liang, Z., Lee, B. W., & Chua, K. Y. (2009). Mite amylase from *Blomia tropicalis* (Blo t 4): differential allergenicity linked to geographical regions. *International archives of allergy and immunology*, 149(1), 25–32. doi:10.1159/000176303
- Chew, F.T., Yi, F. C., Chua, K.Y., Fernandez-Caldas, E., Arruda, L.K., Chapman, M.D., & Lee, B. W. (1999). Allergenic differences between the domestic mites *Blomia tropicalis* and *Dermatophagoides pteronyssinus*. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*, 29(7), 982–8. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10383600>
- Chew, G.L., Reardon, A.M., Correa, J.C., Young, M., Acosta, L., Mellins, R., Chew, F.T., & Perzanowski, M. S. (2009). Mite sensitization among Latina women in New York, where dust-mite allergen levels are typically low. *Indoor air*, 19(3), 193–7. doi:10.1111/j.1600-0668.2008.00578.x
- Chiou, Y.-L., & Lin, C.-Y. (2009). Der p2 activates airway smooth muscle cells in a TLR2/MyD88-dependent manner to induce an inflammatory response. *Journal of cellular physiology*, 220(2), 311–8. doi:10.1002/jcp.21764
- Chou, H., Tam, M.F., Lee, S-S., Tang, R-B., Lin, T-H., Tai, H-Y., Chen, Y-S., & Shen, H-D. (2011). Asp159 is a critical core amino acid of an IgE-binding and cross-reactive epitope of a dust mite allergen Der f 7. *Molecular immunology*, 48(15-16), 2130–4. doi:10.1016/j.molimm.2011.07.004
- Chruszcz, M., Chapman, M. D., Vailes, L. D., Stura, E. a, Saint-Remy, J.-M., Minor, W., & Pomés, A. (2009). Crystal structures of mite allergens Der f 1 and Der p 1 reveal differences in surface-exposed residues that may influence antibody binding. *Journal of molecular biology*, 386(2), 520–30. doi:10.1016/j.jmb.2008.12.049
- Chruszcz, M., Pomés, A., Glesner, J., Vailes, L. D., Osinski, T., Porebski, P. J., Majorek, K. a, Heymann, P.W., Platts-Mills, T.A.E., Minor, W., & Chapman, M. D. (2012). Molecular determinants for antibody binding on group 1 house dust mite allergens. *The Journal of biological chemistry*, 287(10), 7388–98. doi:10.1074/jbc.M111.311159
- Chua, K.Y., Stewart, G.A., Thomas, W.R., Simpson, R.J., Dilworth, R.J., Plozza, T.M., & Turner, K. . (1988). Sequence analysis of cDNA coding for a major house dust mite allergen, Der p 1. Homology with cysteine proteases. *The Journal of experimental medicine*, 167(1), 175–82. Retrieved from <http://jem.rupress.org/content/167/1/175.abstract>
- Chua, K. Y., Dilworth, R. J., & Thomas, W. R. (1990). Expression of *Dermatophagoides pteronyssinus* allergen, Der p II, in *Escherichia coli* and the binding studies with human IgE. *International archives of allergy and applied immunology*, 91(2), 124–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2187817>

- Chua, K. Y., Cheong, N., Kuo, I. C., Lee, B. W., Yi, F. C., Huang, C. H., & Liew, L. N. (2007). The *Blomia tropicalis* allergens. *Protein and peptide letters*, 14(4), 325–33. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17504089>
- Cochrane, G., Akhtar, R., Bonfield, J., Bower, L., Demiralp, F., Faruque, N., Gibson, R., Hoad, G., Hubbard, T., Hunter, C., Jang, M., Juhos, S., Leinonen, R., & Le, E. (2009). Petabyte-scale innovations at the European Nucleotide Archive. *Nucleic acids research*, 37(Database issue), D19–25. doi:10.1093/nar/gkn765
- Collins, S. P., Ball, G., Vonarx, E., Hosking, C., Shelton, M., Hill, D., & Howden, M. E. (1996). Absence of continuous epitopes in the house dust mite major allergens Der p I from *Dermatophagoides pteronyssinus* and Der f I from *Dermatophagoides farinae*. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*, 26(1), 36–42. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8789541>
- Comeau, S. R., Gatchell, D. W., Vajda, S., & Camacho, C. J. (2004). ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, 20(1), 45–50. doi:10.1093/bioinformatics/btg371
- Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., & Kollman, P. A. (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society*, 117(19), 5179–5197. doi:10.1021/ja00124a002
- Cui, Y., Zhou, Y., Ma, G., & Yang, L. (2012). Cloning, bioinformatics analysis, and expression of the dust mite allergen Der f 5 of *Dermatophagoides farinae*. *Brazilian Journal of medical and Biological Research Online Provisional Version, ISSN 0100-*. Retrieved from [http://www.scielo.br/scielo.php?pid=S0100-879X2012007500077&script=sci\\_arttext](http://www.scielo.br/scielo.php?pid=S0100-879X2012007500077&script=sci_arttext)

## **D**

- Da Silva, C. A., Pochard, P., Lee, C. G., & Elias, J. A. (2010). Chitin particles are multifaceted immune adjuvants. *American journal of respiratory and critical care medicine*, 182(12), 1482–91. doi:10.1164/rccm.200912-1877OC
- Dall'Antonia, F., Gieras, A., Devanaboyina, S. C., Valenta, R., & Keller, W. (2011). Prediction of IgE-binding epitopes by means of allergen surface comparison and correlation to cross-reactivity. *The Journal of allergy and clinical immunology*, 128(4), 872–879.e8. doi:10.1016/j.jaci.2011.07.007

- de Halleux, S., Stura, E., VanderElst, L., Carlier, V., Jacquemin, M., & Saint-Remy, J.-M. (2006). Three-dimensional structure and IgE-binding properties of mature fully active Der p 1, a clinically relevant major allergen. *The Journal of allergy and clinical immunology*, *117*(3), 571–6. doi:10.1016/j.jaci.2005.11.032
- DeLano, W. L. (2002). PyMOL: An Open-Source Molecular Graphics Tool. *DeLano Scientific, San Carlos, CA* ([www.pymol.org](http://www.pymol.org)).
- Deb, R., Shakib, F., Reid, K., & Clark, H. (2007). Major house dust mite allergens *Dermatophagoides pteronyssinus* 1 and *Dermatophagoides farinae* 1 degrade and inactivate lung surfactant proteins A and D. *The Journal of biological chemistry*, *282*(51), 36808–19. doi:10.1074/jbc.M702336200
- Dehouck, Y., Kwasigroch, J. M., Gilis, D., & Rooman, M. (2011). PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. *BMC bioinformatics*, *12*(1), 151. doi:10.1186/1471-2105-12-151
- Dehouck, Y., Gilis, D., & Rooman, M. (2012). Design of modified proteins using knowledge-based approaches, *139*(1), 139–147. doi:10.1063/1.4730653
- Derewenda, U., Li, J., & Derewenda, Z. (2002). The crystal structure of a major dust mite allergen Der p 2, and its biological implications. *Journal of molecular biology*, *283*(6)(02), 189–197. doi:00.0000/S0022-2836(02)00027-X
- Dilworth, R. J., Chua, K. Y., & Thomas, W. R. (1991). Sequence analysis of cDNA coding for a major house dust mite allergen, Der f I. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*, *21*(1), 25–32. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2021874>
- Dominguez, C., Boelens, R., & Bonvin, A. M. J. J. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, *125*(7), 1731–7. doi:10.1021/ja026939x
- Douagui, H. (2007). Projet d'application du programme « Global Alliance against Chronic Respiratory Diseases » (GARD) en Algérie. *Revue Française d'Allergologie et d'Immunologie Clinique*, *47*(3), 248–249. doi:10.1016/j.allerg.2007.01.037
- Douwes, J., Zuidhof, a, Doekes, G., van der Zee, S. C., Wouters, I., Boezen, M. H., & Brunekreef, B. (2000). (1-->3)-beta-D-glucan and endotoxin in house dust and peak flow variability in children. *American journal of respiratory and critical care medicine*, *162*(4 Pt 1), 1348–54. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11029343>
- Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., & Liang, J. (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical

mapping of functionally annotated residues. *Nucleic acids research*, 34(Web Server issue), W116–8. doi:10.1093/nar/gkl282

Durham, S. R. (2006). Allergen immunotherapy (desensitisation) for allergic diseases, 6(4), 4–7.

## **E**

ECRHS-I, The European Community Respiratory Health Survey I Steering Committee (1996). Variations in the prevalence of respiratory symptoms, self-reported asthma attacks, and use of asthma medication in the European Community Respiratory Health Survey (ECRHS). *European Respiratory Journal*, 9(4), 687–695. doi:10.1183/09031936.96.09040687

ECRHS-II, The European Community Respiratory Health Survey II Steering Committee. (2002). The European Community Respiratory Health Survey II. *European Respiratory Journal*, 20(5), 1071–1079. doi:10.1183/09031936.02.00046802

Efremov, R. G., & Alix, a J. (1993). Environmental characteristics of residues in proteins: three-dimensional molecular hydrophobicity potential approach. *Journal of biomolecular structure & dynamics*, 11(3), 483–507. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8129869>

El-Manzalawy, Y., Dobbs, D., & Honavar, V. (2008). Predicting linear B-cell epitopes using string kernels. *Journal of molecular recognition: JMR*, 21(4), 243–55. doi:10.1002/jmr.893

El-Manzalawy, Y., & Honavar, V. (2010). Recent advances in B-cell epitope prediction methods. *Immunome research*, 6 Suppl 2(Suppl 2), S2. doi:10.1186/1745-7580-6-S2-S2

Elsayed, S., & Aas, K. (1970). Characterization of a major allergen (cod). Chemical composition and immunological properties. *International archives of allergy and immunology*, 38(5), 536–48. Retrieved from <http://content.karger.com/ProdukteDB/produkte.asp?Doi=230307>

Epton, M. J., Dilworth, R. J., Smith, W., & Thomas, W. R. (2001). Sensitisation to the lipid-binding apolipoprotein allergen Der p 14 and the peptide Mag-1. *International archives of allergy and immunology*, 124(1-3), 57–60. doi:53668

Eriksson, T.L., Johansson, E., Whitley, P., Schmidt, M., Elsayed, S., & van Hage-Hamsten, M. (1998). Cloning and characterisation of a group II allergen from the dust mite *Tyrophagus putrescentiae*. *European journal of biochemistry / FEBS*, 251(1-2), 443–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9492316>

Eriksson, T.L., Rasool, O., Huecas, S., Whitley, P., Cramer, R., Appenzeller, U., Gafvelin, G., & van Hage-Hamsten, M. (2001). Cloning of three new allergens from the dust mite *Lepidoglyphus destructor* using phage surface display technology. *European journal of biochemistry / FEBS*, 268(2), 287–94. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11168362>

Ezkurdia, I., Bartoli, L., Fariselli, P., Casadio, R., Valencia, A., & Tress, M. L. (2009). Progress and challenges in predicting protein-protein interaction sites. *Briefings in bioinformatics*, 10(3), 233–46. doi:10.1093/bib/bbp021

## F

Fahlbusch, B., Koch, A., Douwes, J., Bischof, W., Gehring, U., Richter, K., Wichmann, H-E., & Heinrich, J. (2003). The effect of storage on allergen and microbial agent levels in frozen house dust. *Allergy*, 58(2), 150–3. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12622747>

Fernández-Caldas, E. (2002). Dust mite allergens: mitigation and control. *Current allergy and asthma reports*, 12165210. Retrieved from <http://www.springerlink.com/index/575313377VG665X1.pdf>

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., & Scheiner, O. (1998). Modulation of IgE reactivity of allergens by site-directed mutagenesis: potential use of hypoallergenic variants for immunotherapy. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 12(2), 231–42. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9472988>

Flicker, S., Steinberger, P., Ball, T., Krauth, M-T., Verdino, P., Valent, P., Almo, S., & Valenta, R. (2006). Spatial clustering of the IgE epitopes on the major timothy grass pollen allergen Phl p 1: importance for allergenic activity. *The Journal of allergy and clinical immunology*, 117(6), 1336–43. doi:10.1016/j.jaci.2006.02.012

Focke, M., Mahler, V., Ball, T., Sperr, W.R., Majlesi, Y., Valent, P., Kraft, D., & Valenta, R. (2001). Nonanaphylactic synthetic peptides derived from B cell epitopes of the major grass pollen allergen, Phl p 1, for allergy vaccination. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 15(11), 2042–4. doi:10.1096/fj.01-0016fje

Folgori, A., Tafi, R., Meola, A., Felici, F., Galfré, G., Cortese, R., Monaci, P., & Nicosia, A. (1994). A general strategy to identify mimotopes of pathological antigens using only random peptide libraries and human sera. *The EMBO journal*, 13(9), 2236–43. Retrieved from

<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=395079&tool=pmcentrez&rendertype=abstract>

- Friedland, N., Liou, H.-L., Lobel, P., & Stock, A. M. (2003). Structure of a cholesterol-binding protein deficient in Niemann-Pick type C2 disease. *Proceedings of the National Academy of Sciences of the United States of America*, 100(5), 2512–7. doi:10.1073/pnas.0437840100
- Fuiano, N., & Incorvaia, C. (2012). Dissecting the causes of atopic dermatitis in children: less foods, more mites. *Allergology international : official journal of the Japanese Society of Allergology*, 61(2), 231–43. doi:10.2332/allergolint.11-RA-0371
- Fujikawa, A., Ishimaru, N., Seto, A., Yamada, H., Aki, T., Shigeta, S., Wada, T., Jyo, T., Murooka, Y., Oka, S., & Ono, K. (1996). Cloning and characterization of a new allergen, Mag 3, from the house dust mite, *Dermatophagoides farinae*: cross-reactivity with high- molecular-weight allergen. *Molecular immunology*, 33(3), 311–319.
- Furmonaviciene, R., Ghaemmaghami, A.M., Boyd, S.E., Jones, N.S., Bailey, K., Willis, A.C., Sewell, H.F., Mitchell, D.A., & Shakib, F. (2007). The protease allergen Der p 1 cleaves cell surface DC-SIGN and DC-SIGNR: experimental analysis of in silico substrate identification and implications in allergic responses. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*, 37(2), 231–42. doi:10.1111/j.1365-2222.2007.02651.x

## G

- Gafvelin, G., Johansson, E., Lundin, A., Smith, A.M., Chapman, M.D., Benjamin, D.C., Derewenda U., & Van Hage-Hamsten, M. (2001). Cross-reactivity studies of a new group 2 allergen from the dust mite *Glycyphagus domesticus*, Gly d 2, and group 2 allergens from *Dermatophagoides pteronyssinus*, *Lepidoglyphus destructor*, and *Tyrophagus putrescentiae* with recombinant allergens. *The Journal of allergy and clinical immunology*, 107(3), 511–8. doi:10.1067/mai.2001.112264
- Galli, S. J., Tsai, M., & Piliponsky, A. M. (2008). The development of allergic inflammation. *Nature*, 454(7203), 445–54. doi:10.1038/nature07204
- Gao, Y. F., Wang, D. Y., Ong, T. C., Tay, S. L., Yap, K. H., & Chew, F. T. (2007). Identification and characterization of a novel allergen from *Blomia tropicalis*: Blo t 21. *The Journal of allergy and clinical immunology*, 120(1), 105–12. doi:10.1016/j.jaci.2007.02.032
- Ghaemmaghami, A., Gough, L., Sewell, H. F., & Shakib, F. (2002). The proteolytic activity of the major dust mite allergen Der p 1 conditions dendritic cells to produce less interleukin-12: allergen-induced Th2 bias determined at the dendritic cell level. *Clinical*

*and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*, 32(10), 1468–75. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12372127>

Ghosh, D., & Gupta-Bhattacharya, S. (2011). Allergen Bioinformatics: Recent Trends and Developments. *Selected Works in Bioinformatics*, 1–17. doi:10.5772/20269

Gracy, J., Chiche, L., & Sallantin, J. (1993). Improved alignment of weakly homologous protein sequences using structural information. *Protein engineering*, 6(8), 821–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8309929>

Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A., & Baker, D. (2003). Protein–Protein Docking with Simultaneous Optimization of Rigid-body Displacement and Side-chain Conformations. *Journal of Molecular Biology*, 331(1), 281–299. doi:10.1016/S0022-2836(03)00670-3

Greene, W. K., & Thomas, W. R. (1992). IgE binding structures of the major house dust mite allergen Der p I. *Molecular immunology*, 29(2), 257–62. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1371823>

Gregory, L. G., & Lloyd, C. M. (2011). Orchestrating house dust mite-associated allergy in the lung. *Trends in immunology*, 32(9), 402–11. doi:10.1016/j.it.2011.06.006

Grosdidier, A., Zoete, V., & Michielin, O. (2007). EADock: Docking of Small Molecules Into Protein Active Sites With a Multiobjective Evolutionary Optimization. *PROTEINS: Structure, Function, and Bioinformatics*, 67, 1010–1025. doi:10.1002/prot

## H

Hakkaart, G. A., Chapman, M. D., Aalberse, R. C., & van Ree, R. (1998). Immune-reactivity of recombinant isoforms of the major house dust mite allergen Der p 2. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*, 28(2), 169–74. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9515589>

Hales, B.J., Martin, A.C., Pearce, L.J., Laing, I.A., Hayden, C.M., Goldblatt, J., Le Souëf, P.N., & Thomas, W. R. (2006). IgE and IgG anti-house dust mite specificities in allergic disease. *The Journal of allergy and clinical immunology*, 118(2), 361–7. doi:10.1016/j.jaci.2006.04.001

Hantusch, B., Krieger, S., Untermayr, E., Schöll, I., Knittelfelder, R., Flicker, S., Spitzauer, S., Valenta, R., Boltz-Nitulescu, G., Scheiner, O., & Jensen-Jarolim, E. (2004). Mapping of conformational IgE epitopes on Phl p 5a by using mimotopes from a phage

- display library. *The Journal of allergy and clinical immunology*, 114(6), 1294–300. doi:10.1016/j.jaci.2004.06.048
- Henikoff, S., & Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences of the United States of America*, 89(22), 10915–9. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=50453&tool=pmcentrez&rendertype=abstract>
- Hileman, R.E., Silvanovich, A., Goodman, R.E., Rice, E.A., Holleschak, G., Astwood, J.D., & Hefle, S. L. (2002). Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. *International archives of allergy and immunology*, 128(4), 280–91. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12218366>
- Hopp, T. P., & Woods, K. R. (1981). Prediction of protein antigenic determinants from amino acid sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 78(6), 3824–8. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=319665&tool=pmcentrez&rendertype=abstract>
- Hopp, T. P. (1993). Retrospective: 12 years of antigenic determinant predictions, and more. *Peptide research*, 6(4), 183–90. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7691280>
- Huang, J., & Honda, W. (2006). CED: a conformational epitope database. *BMC immunology*, 7, 7. doi:10.1186/1471-2172-7-7
- Huang, Y. X., Bao, Y. L., Guo, S. Y., Wang, Y., Zhou, C. G., & Li, Y. X. (2008). Pep-3D-Search: a method for B-cell epitope prediction based on mimotope analysis. *BMC bioinformatics*, 9, 538. doi:10.1186/1471-2105-9-538
- Huang, C.-H., Loo, E. X.-L., Kuo, I.-C., Soh, G. H., Goh, D. L.-M., Lee, B. W., & Chua, K. Y. (2011). Airway inflammation and IgE production induced by dust mite allergen-specific memory/effector Th2 cell line can be effectively attenuated by IL-35. *Journal of immunology (Baltimore, Md. : 1950)*, 187(1), 462–71. doi:10.4049/jimmunol.1100259
- Huang, J., Ru, B., Zhu, P., Nie, F., Yang, J., Wang, X., Dai, P., Lin, H., Guo, F.-B., & Rao, N. (2012). MimoDB 2.0: a mimotope database and beyond. *Nucleic acids research*, 40(Database issue), D271–7. doi:10.1093/nar/gkr922

**I**

Ichikawa, S., Hatanaka, H., Yuuki, T., Iwamoto, N., Kojima, S., Nishiyama, C., Ogura, K., Okumura, Y., & Inagaki, F. (1998). Solution structure of Der f 2, the major mite allergen for atopic diseases. *The Journal of biological chemistry*, 273(1), 356–60. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9417088>

Ichikawa, S., Takai, T., Yashiki, T., Takahashi, S., Okumura, K., Ogawa, H., Kohda, D., & Hatanaka, H. (2009). Lipopolysaccharide binding of the mite allergen Der f 2. *Genes to cells: devoted to molecular & cellular mechanisms*, 14(9), 1055–65. doi:10.1111/j.1365-2443.2009.01334.x

Ipsen, H., Henmar, H., Bolwing, C., Duffort, O., Barber, D., Polo, F., & Larsen, J. (2004). Mapping of Der p 2 antibody binding epitopes by site directed mutagenesis. *Journal of Allergy and Clinical Immunology*, 113(2), S144. doi:10.1016/j.jaci.2003.12.520

ISAAC, The International Study of Asthma and Allergies in Childhood (ISSAC) Steering Committee. (1998a). Worldwide variations in the prevalence of asthma symptoms: the International Study of Asthma and Allergies in Childhood (ISAAC). *European Respiratory Journal*, 12(2), 315–335. doi:10.1183/09031936.98.12020315

ISAAC, The International Study of Asthma and Allergies in Childhood (ISAAC) Steering committee. (1998b). Worldwide variation in prevalence of symptoms of asthma , allergic rhinoconjunctivitis , and atopic eczema : ISAAC. *The Lancet*, 351, 1225–1232. Retrieved from [http://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(97\)07302-9](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(97)07302-9)

Ivanciuc, O., Schein, C. H., & Braun, W. (2003). SDAP: database and computational tools for allergenic proteins. *Nucleic Acids Research*, 31(1), 359–362. doi:10.1093/nar/gkg010

**J**

Jacquet, A. (2009). New Insights into the Molecular Basis of the House Dust Mite-Induced Allergic Response. *The Open Allergy Journal*, 2(1), 38–44. doi:10.2174/1874838400902010038

Jacquet, A. (2011a). The role of the house dust mite-induced innate immunity in development of allergic response. *International archives of allergy and immunology*, 155(2), 95–105. doi:10.1159/000320375

Jacquet, A. (2011b). The role of innate immunity activation in house dust mite allergy. *Trends in molecular medicine*, 17(10), 604–11. doi:10.1016/j.molmed.2011.05.014

- Jacquet, A. (2011c). Interactions of airway epithelium with protease allergens in the allergic response. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*, 41(3), 305–11. doi:10.1111/j.1365-2222.2010.03661.x
- Jauch, R., Yeo, H. C., Kolatkar, P. R., & Clarke, N. D. (2007). Assessment of CASP7 structure predictions for template free targets. *Proteins*, 69 Suppl 8(c), 57–67. doi:10.1002/prot.21771
- Jeannin, P., Didierlaurent, A., Gras-Masse, H., Ellass, A.A., Delneste, Y., Cardot, E., Joseph, M., Tartar, A., Vergoten, G., & Pestel, J. (1992). Specific histamine release capacity of peptides selected from the modeled Der p I protein, a major allergen of *Dermatophagoides pteronyssinus*. *Molecular immunology*, 29(6), 739–49. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1376413>
- Jeong, K., Kim, W., Lee, J., & Lee, J. (2005). Immunoglobulin E reactivity of recombinant allergen Tyr p 13 from *Tyrophagus putrescentiae* homologous to fatty acid binding protein. *Clinical and Diagnostic Laboratory Immunology*, 12(5), 581–585. doi:10.1128/CDLI.12.5.581
- Jeong, K. Y., Lee, H., Lee, J. S., Lee, J., Lee, I-Y., Ree, H-I., Hong, C-S., & Yong, T.-S. (2007). Molecular cloning and the allergenic characterization of tropomyosin from *Tyrophagus putrescentiae*. *Protein and peptide letters*, 14(5), 431–6. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17584167>
- Jeong, K., Kim, C., Un, S., & Yi, M. (2010). Allergenicity of Recombinant Troponin C from *Tyrophagus putrescentiae*. *International Archives of Allergy and Immunology*, 151(3), 207–13. Retrieved from <http://content.karger.com/ProdukteDB/produkte.asp?Doi=242358>
- Jeong, K. Y., Park, J.-W., & Hong, C.-S. (2012). House dust mite allergy in Korea: the most important inhalant allergen in current and future. *Allergy, asthma & immunology research*, 4(6), 313–25. doi:10.4168/aaair.2012.4.6.313
- Johansson, S., Hourihane, J., Bousquet, P.-J., Bruijnzeel-Koomen, C., Dreborg, S., Haahtela, T., Kowalski, M., Myging, N., Ring, J., van Cauwenberge, P., & Wüthrich, B. (2001). A revised nomenclature for allergy: an EAACI position statement from the EAACI nomenclature task force. *Allergy*, 56, 813–824. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1398-9995.2001.00002.x-i1/full>
- Jones, G., Willett, P., Glen, R.C., Leach, A.R., & Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *Journal of molecular biology*, 267(3), 727–48. doi:10.1006/jmbi.1996.0897

Jones, D. T. (1998). THREADER: Protein Sequence Threading by Double Dynamic Programming. (in) Computational Methods in Molecular Biology. Steven Salzberg, David Searls, and Simon Kasif, Eds. Elsevier Science. Chapter 13.

Joseph-McCarthy, D. (1999). Computational approaches to structure-based ligand design. *Pharmacology & therapeutics*, 84(2), 179–91. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/21554079>

## K

Kabsch, W. (1978). A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, 34(5), 827–828. doi:10.1107/S0567739478001680

Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 12, 2577–637. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/bip.360221211/abstract>

Kaiser, L., Velickovic, T.C., Badia-Martinez, D., Adedoyin, J., Thunberg, S., Hallén, D., Berndt, K., Grönlund, H., Gafvelin, G., van Hage, M., & Achour, A. (2007). Structural characterization of the tetrameric form of the major cat allergen Fel d 1. *Journal of molecular biology*, 370(4), 714–27. doi:10.1016/j.jmb.2007.04.074

Kalsheker, N. a, Deam, S., Chambers, L., Sreedharan, S., Brocklehurst, K., & Lomas, D. a. (1996). The house dust mite allergen Der p1 catalytically inactivates alpha 1-antitrypsin by specific reactive centre loop cleavage: a mechanism that promotes airway inflammation and asthma. *Biochemical and biophysical research communications*, 221(1), 59–61. doi:10.1006/bbrc.1996.0544

Kaminuma, E., Kosuge, T., Kodama, Y., Aono, H., Mashima, J., Gojobori, T., Sugawara, H., Ogasawara, O., Takagi, T., Okubo, K., & Nakamura, Y. (2011). DDBJ progress report. *Nucleic acids research*, 39(Database issue), D22–7. doi:10.1093/nar/gkq1041

Karplus, P., & Schulz, G. (1985). Prediction of chain flexibility in proteins. *Naturwissenschaften*, 72, 2–3. Retrieved from <http://www.springerlink.com/index/Q1224Q1555013J53.pdf>

Kauffman, H.F., Tamm, M., Timmerman, J.A.B., & Borger, P. (2006). House dust mite major allergens Der p 1 and Der p 5 activate human airway-derived epithelial cells by protease-dependent and protease-independent mechanisms. *Clinical and molecular allergy : CMA*, 4, 5. doi:10.1186/1476-7961-4-5.

Kawamoto, S., Mizuguchi, Y., Morimoto, K., Aki, T., Shigeta, S., Yasueda, H., Wada, T., Suzuki, O., Jyo, T., & Ono, K. (1999). Cloning and expression of Der f 6, a serine

- protease allergen from the house dust mite, *Dermatophagoides farinae*. *Biochimica et biophysica acta*, 1454(2), 201–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10381565>
- Kawamoto, S., Suzuki, T., Aki, T., Katsutani, T., & Tsuboi, S. (2002). Der f 16 : a novel gelsolin-related molecule identified as an allergen from the house dust mite , *Dermatophagoides farinae*. *FEBS letters*, 516, 234–238.
- Kelley, L. A., MacCallum, R. M., & Sternberg, M. J. E. (1999). Recognition of remote protein homologies using three-dimensional information to generate a position specific scoring matrix in the program 3D-PSSM. *Proceedings of the third annual international conference on Computational molecular biology - RECOMB '99*, 218–225. doi:10.1145/299432.299486
- Kelley, L. A., & Sternberg, M. J. E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols*, 4(3), 363–71. doi:10.1038/nprot.2009.2
- Khemili, S., Kwasigroch, J. M., Hamadouche, T., & Gilis, D. (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. *Journal of biomolecular structure & dynamics*, 29(4), 663–75. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/22208271>
- Kidon, M.I., Chiang, W.C., Liew, W.K., Kang, Ong T.C., Tiong, Y.S., Wong, K.N., Aaron, C.A, Ong, S.T., Gao, Y.F., Kavita, R., Bi, X.Z., Shang H?S., & Chew, F. T. (2011). Mite component-specific IgE repertoire and phenotypes of allergic disease in childhood: the tropical perspective. *Pediatric allergy and immunology : official publication of the European Society of Pediatric Allergy and Immunology*, 22(2), 202–10. doi:10.1111/j.1399-3038.2010.01094.x
- Kim, S., Lee, S. H., & Park, D. (2001). Leucine zipper-mediated homodimerization of the p21-activated kinase-interacting factor, beta Pix. Implication for a role in cytoskeletal reorganization. *The Journal of biological chemistry*, 276(14), 10581–4. doi:10.1074/jbc.C000806200
- Kim, H. M., Park, B. S., Kim, J.-I., Kim, S. E., Lee, J., Oh, S. C., Enkhbayar, P., Matsushima, N., Lee, H., Yoo, O.J., & Lee, J.-O. (2007). Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell*, 130(5), 906–17. doi:10.1016/j.cell.2007.08.002
- King, C., Simpson, R. J., Moritz, R. L., Reed, G. E., Thompson, P. J., & Stewart, G. A. (1996). The isolation and characterization of a novel collagenolytic serine protease allergen (Der p 9) from the dust mite *Dermatophagoides pteronyssinus*. *The Journal of allergy and clinical immunology*, 98(4), 739–47. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8876548>

- King, T. P., & Spangfort, M. D. (2000). Structure and biology of stinging insect venom allergens. *International archives of allergy and immunology*, 123(2), 99–106. doi:24440
- Kozakov, D., Brenke, R., Comeau, S. R., & Vajda, S. (2006). PIPER: An FFT-based protein docking program with pairwise potentials. *PROTEINS: Structure, Function, and Bioinformatics*, 65, 392–406. doi:10.1002/prot
- Kozakov, D., Hall, D. R., Beglov, D., Brenke, R., Comeau, S.R., Shen, Y., Li, K., Zheng, J., Vakili, P., Paschalidis, I.C., & Vajda, S. (2010). Achieving reliability and high accuracy in automated protein docking: ClusPro, PIPER, SDU, and stability analysis in CAPRI rounds 13-19. *Proteins*, 78(15), 3124–30. doi:10.1002/prot.22835
- Kramer, B., Rarey, M., & Lengauer, T. (1999). Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. *Proteins*, 37(2), 228–41. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10584068>
- Kronqvist, M., Johansson, E., Pershagen, G., Johansson, S. G., & van Hage-Hamsten, M. (1999). Increasing prevalence of asthma over 12 years among dairy farmers on Gotland, Sweden: storage mites remain dominant allergens. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*, 29(1), 35–41. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10051700>
- Kulikova, T., Aldebert, P., Althorpe, N., Baker, W., Bates, K., Browne, P., van den Broek, A., Cochrane, G., Duggan, K., Eberhardt, R., Faruque, N., Garcia-Pastor, M., & Harte, N. R. (2004). The EMBL Nucleotide Sequence Database. *Nucleic acids research*, 32(Database issue), D27–30. doi:10.1093/nar/gkh120
- Kulkarni-Kale, U., Bhosle, S., & Kolaskar, A. S. (2005). CEP: a conformational epitope prediction server. *Nucleic acids research*, 33(Web Server issue), W168–71. doi:10.1093/nar/gki460
- Kundu, S., & Roy, D. (2010). Structural study of biologically significant ligands with major birch pollen allergen Betv1 by docking and molecular dynamics simulation. *Bioinformation*, 2063(7), 326–330. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2957764/>
- Kuo, I.C., Cheong, N., Shek, L.P.C., Yi, F.C., Chew, F.T., Lee, B.W., & Chua, K. (1999). Comparative Study between Singapore and Taiwan, 179–188.
- Kuo, I. C., Cheong, N., Trakultivakorn, M., Lee, B. W., & Chua, K. Y. (2003). An extensive study of human IgE cross-reactivity of Blo t 5 and Der p 5. *Journal of Allergy and Clinical Immunology*, 111(3), 603–609. doi:10.1067/mai.2003.167

£

- Lake, F., Ward, L., & Simpson, R. (1991). House dust mite-derived amylase: Allergenicity and physicochemical characterization. *Journal of Allergy and Clinical Immunology*, 87(6), 1035–42. Retrieved from <http://www.sciencedirect.com/science/article/pii/009167499192147S>
- Larché, M., Akdis, C. A., & Valenta, R. (2006). Immunological mechanisms of allergen-specific immunotherapy. *Nature reviews. Immunology*, 6(10), 761–71. doi:10.1038/nri1934
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)*, 23(21), 2947–8. doi:10.1093/bioinformatics/btm404
- Larsen, J. E. P., Lund, O., & Nielsen, M. (2006). Improved method for predicting linear B-cell epitopes. *Immunome research*, 2, 2. doi:10.1186/1745-7580-2-2
- Lascombe, M.B., Grégoire, C., Poncet, P., Tavares, G.A., Rosinski-Chupin, I., Rabillon, J., Goubran-Botros, H., Mazié, J. C., David, B., & Alzari, P. M. (2000). Crystal structure of the allergen Equ c 1. A dimeric lipocalin with restricted IgE-reactive epitopes. *The Journal of biological chemistry*, 275(28), 21572–7. doi:10.1074/jbc.M002854200
- Laskowski, R. a., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283–291. doi:10.1107/S0021889892009944
- Lee, C-S., Tsai, L-C., Chao, P-L., Lin, C-Y., Hung, M-W., Chien, A-I., Chiang, Y-T., & Han, S-H. (2004). Protein sequence analysis of a novel 103-kDa *Dermatophagoides pteronyssinus* mite allergen and prevalence of serum immunoglobulin E reactivity to rDer p 11 in allergic adult patients. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*, 34(3), 354–62. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15005727>
- Lesk, V. I., & Sternberg, M. J. E. (2008). 3D-Garden: a system for modelling protein-protein complexes based on conformational refinement of ensembles generated with the marching cubes algorithm. *Bioinformatics (Oxford, England)*, 24(9), 1137–44. doi:10.1093/bioinformatics/btn093
- Levitt, M. (1992). Accurate Modeling of Protein Conformation by Automatic Segment Matching. *Journal of molecular biology*, 266, 407–533.

- Lewkowich, I.P., Herman, N.S., Schleifer, K.W., Dance, M.P., Chen, B.L., Dienger, K.M., Sproles, A.A., Shah, J.S., Köhl, J., Belkaid, Y., & Wills-Karp, M. (2005). CD4+CD25+ T cells protect against experimentally induced asthma and alter pulmonary dendritic cell phenotype and function. *The Journal of experimental medicine*, 202(11), 1549–61. doi:10.1084/jem.20051506
- Liang, S., Zheng, D., Zhang, C., & Zacharias, M. (2009). Prediction of antigenic epitopes on protein surfaces by consensus scoring. *BMC bioinformatics*, 10, 302. doi:10.1186/1471-2105-10-302
- Liang, S., Zheng, D., Standley, D. M., Yao, B., Zacharias, M., & Zhang, C. (2010). EPSVR and EPMeta: prediction of antigenic epitopes using support vector regression and multiple server results. *BMC bioinformatics*, 11, 381. doi:10.1186/1471-2105-11-381
- Liao, E-C., Hsu, E-L., Tsai, J-J., & Ho, C.-M. (2009). Immunologic characterization and allergenicity of recombinant Tyr p 3 allergen from the storage mite *Tyrophagus putrescentiae*. *International archives of allergy and immunology*, 150(1), 15–24. doi:10.1159/000210376
- Lintner, T. J., & Brame, K. A. (1993). The effects of season, climate, and air-conditioning on the prevalence of *Dermatophagoides* mite allergens in household dust. *The Journal of allergy and clinical immunology*, 91(4), 862–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8473674>
- Linhart, B., & Valenta, R. (2012). Vaccines for allergy. *Current opinion in immunology*, 24(3), 354–60. doi:10.1016/j.coi.2012.03.006
- Liu, R., & Hu, J. (2011). Prediction of Discontinuous B-Cell Epitopes Using Logistic Regression and Structural Information. *Journal of Proteomics & Bioinformatics*, 04(01), 10–15. doi:10.4172/jpb.1000161
- Lynch, N. R., Thomas, W. R., Garcia, N. M., Di Prisco, M. C., Puccio, F. A., L'opez, R. I., Hazell, L. A., & Shen, H. D., Lin, K. L., & Chua, K. Y. (1997). Biological activity of recombinant Der p 2, Der p 5 and Der p 7 allergens of the house-dust mite *Dermatophagoides pteronyssinus*. *International archives of allergy and immunology*, 114(1), 59–67. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9303332>
- Lynch, N. R., Puccio, F. a, Di Prisco, M. C., Lopez, R. I., Hazell, L. a, Smith, W. a, & Thomas, W. R. (1998). Reactivity to recombinant house-dust-mite allergens in asthma and rhinitis in a tropical environment. *Allergy*, 53(8), 808–11. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9722232>

**M**

- Machado, D. C., Horton, D., Harrop, R., Peachell, P. T., & Helm, B. A. (1996). Potential allergens stimulate the release of mediators of the allergic response from cells of mast cell lineage in the absence of sensitization with antigen-specific IgE. *European journal of immunology*, 26(12), 2972–80. doi:10.1002/eji.1830261224
- Mari, A., Scala, E., Palazzo, P., Ridolfi, S., Zennaro, D., & Carabella, G. (2006). Bioinformatics applied to allergy: allergen databases, from collecting sequence information to data integration. The Allergome platform as a model. *Cellular immunology*, 244(2), 97–100. doi:10.1016/j.cellimm.2007.02.012
- Maruo, K., Akaike, T., Ono, T., Okamoto, T., & Maeda, H. (1997). Generation of anaphylatoxins through proteolytic processing of C3 and C5 by house dust mite protease. *The Journal of allergy and clinical immunology*, 100(2), 253–60. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9275149>
- Mayrose, I., Shlomi, T., Rubinstein, N. D., Gershoni, J. M., Ruppín, E., Sharan, R., & Pupko, T. (2007). Epitope mapping using combinatorial phage-display libraries: a graph-based algorithm. *Nucleic acids research*, 35(1), 69–78. doi:10.1093/nar/gkl975
- McCall, C., Hunter, S., Stedman, K., Weber, E., Hillier, A., Bozic, C., Rivoire, B., & Olivry, T. (2001). Characterization and cloning of a major high molecular weight house dust mite allergen (Der f 15) for dogs. *Veterinary immunology and immunopathology*, 78(3-4), 231–47. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11292526>
- Melo, F., & Feytmans, E. (1998). Assessing protein structures with a non-local atomic interaction energy. *Journal of molecular biology*, 277(5), 1141–52. doi:10.1006/jmbi.1998.1665
- Meloën, R. H., Puijk, W. C., & Slootstra, J. W. (2000). Mimotopes: realization of an unlikely concept. *Journal of molecular recognition : JMR*, 13(6), 352–9. doi:10.1002/1099-1352(200011/12)13:6<352::AID-JMR509>3.0.CO;2-C
- Mills, K. L., Hart, B. J., Lynch, N. R., Thomas, W. R., & Smith, W. (1999). Molecular characterization of the group 4 house dust mite allergen from *Dermatophagoides pteronyssinus* and its amylase homologue from *Euroglyphus maynei*. *International archives of allergy and immunology*, 120(2), 100–7. doi:24227
- Mora, C., Flores, I., Montealegre, F., & Díaz, a. (2003). Cloning and expression of Blo t 1, a novel allergen from the dust mite *Blomia tropicalis*, homologous to cysteine proteases. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*, 33(1), 28–34. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12534546>

- Moreau, V., Fleury, C., Piquer, D., Nguyen, C., Novali, N., Villard, S., Laune, D., Granier, C., & Molina, F. (2008). PEPOP: computational design of immunogenic peptides. *BMC bioinformatics*, 9, 71. doi:10.1186/1471-2105-9-71
- Morgan, M., & Arlian, L. (1997). Characterization of the allergens of the house dust mite *Euroglyphus maynei*. *Journal of Allergy and Clinical Immunology*, 10, 222–228. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0091674997702288>
- Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., & Olson, A. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of computational biology: a journal of computational molecular cell biology*, 19(14), 1639–1662. Retrieved from [http://hidra1.cbm.uam.es/publications/teaching/curso\\_verano\\_uah\\_2011/autodock4.pdf](http://hidra1.cbm.uam.es/publications/teaching/curso_verano_uah_2011/autodock4.pdf)
- Moult, J., Fidelis, K., Kryshtafovych, A., & Tramontano, A. (2011). Critical assessment of methods of protein structure prediction (CASP)--round IX. *Proteins*, 79 Suppl 1, 1–5. doi:10.1002/prot.23200
- Mueller, G. A., Smith, A. M., Chapman, M. D., Rule, G. S., & Benjamin, D. C. (2001). Hydrogen exchange nuclear magnetic resonance spectroscopy mapping of antibody epitopes on the house dust mite allergen Der p 2. *The Journal of biological chemistry*, 276(12), 9359–65. doi:10.1074/jbc.M010812200
- Mueller, G. A., Gosavi, R. A., Krahn, J. M., Edwards, L. L., Cuneo, M. J., Glesner, J., Pomés, A., Chapman, M. D., London, R. E., & Pedersen, L. C. (2010a). Der p 5 crystal structure provides insight into the group 5 dust mite allergens. *The Journal of biological chemistry*, 285(33), 25394–401. doi:10.1074/jbc.M110.128306
- Mueller, G. A., Edwards, L. L., Aloor, J. J., Fessler, M. B., Glesner, J., Pomés, A., Chapman, M. D., London, R. E., & Pedersen, L. C. (2010b). The structure of the dust mite allergen Der p 7 reveals similarities to innate immune proteins. *The Journal of allergy and clinical immunology*, 125(4), 909–917.e4. doi:10.1016/j.jaci.2009.12.016
- Muh, H. C., Tong, J. C., & Tammi, M. T. (2009). AllerHunter: a SVM-pairwise system for assessment of allergenicity and allergic cross-reactivity in proteins. *PloS one*, 4(6), e5861. doi:10.1371/journal.pone.0005861
- Mumey, B. M., Bailey, B. W., Kirkpatrick, B., Jesaitis, A. J., Angel, T., & Dratz, E. A. (2003). A new method for mapping discontinuous antibody epitopes to reveal structural features of proteins. *Journal of computational biology: a journal of computational molecular cell biology*, 10(3-4), 555–67. doi:10.1089/10665270360688183
- Mumey, B., Ohler, N., Angel, T., Jesaitis, A., Dratz, E. (2006). Filtering Epitope Alignments to Improve Protein Surface Prediction. (G. Min, B. Martino, L. T. Yang, M. Guo, & G.

Rünger, Eds.) *Lecture Notes in Computer Science*, 4331, 648–657.  
doi:10.1007/11942634

## N

Nadchatram, M. (2005). House dust mites, our intimate associates. *Tropical biomedicine*, 22(1), 23–37. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16880751>

Naik, M. T., Chang, C.-F., Kuo, I.-C., Kung, C. C.-H., Yi, F.-C., Chua, K.-Y., & Huang, T.-H. (2008). Roles of structure and structural dynamics in the antibody recognition of the allergen proteins: an NMR study on *Blomia tropicalis* major allergen. *Structure (London, England : 1993)*, 16(1), 125–36. doi:10.1016/j.str.2007.10.022

Nakamura, R., Nakamura, R., & Teshima, R. (2009). Major revision of the allergen database for food safety (ADFS) and validation of the motif-based allergenicity prediction tool. *Kokuritsu Iyakuhin Shokuhin Eisei Kenkyūjo hōkoku = Bulletin of National Institute of Health Sciences*, (127), 44–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/20306706>

Nelson, R. P., DiNicolo, R., Fernández-Caldas, E., Seleznick, M. J., Lockey, R. F., & Good, R. a. (1996). Allergen-specific IgE levels and mite allergen exposure in children with acute asthma first seen in an emergency department and in nonasthmatic control subjects. *The Journal of allergy and clinical immunology*, 98(2), 258–63. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8757201>

## O

Odorico, M., & Pellequer, J.-L. (2003). BEPITOPE: predicting the location of continuous epitopes and patterns in proteins. *Journal of molecular recognition : JMR*, 16(1), 20–2. doi:10.1002/jmr.602

Oezguen, N., Zhou, B., Negi, S. S., Ivanciuc, O., Schein, C. H., Labesse, G., Braun, W., Jay, J., & Drive, H. (2009). Comparative 3D-modeling of allergenic proteins and amino acid composition of potential conformational IgE epitopes, 45(14), 3740–3747. doi:10.1016/j.molimm.2008.05.026. ComprehensiveOMS, L'Organisation Mondiale de la santé (2012). <http://www.who.int/>

O'Neil, S. E., Heinrich, T. K., Hales, B. J., Hazell, L. a, Holt, D. C., Fischer, K., & Thomas, W. R. (2006). The chitinase allergens Der p 15 and Der p 18 from *Dermatophagoides pteronyssinus*. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*, 36(6), 831–9. doi:10.1111/j.1365-2222.2006.02497.x

O'Neill, G. M., Donovan, G. R., & Baldo, B. A. (1995). Glutathione S-transferase a major allergen of the house dust mite, *Dermatophagoides pteronyssinus*. *Immunology letters*, 48(2), 103–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8719107>

Ouyang, W., Kolls, J. K., & Zheng, Y. (2008). The biological functions of T helper 17 cell effector cytokines in inflammation. *Immunity*, 28(4), 454–67. doi:10.1016/j.immuni.2008.03.004

## **P**

Park, J. H., Gold, D. R., Spiegelman, D. L., Burge, H. a, & Milton, D. K. (2001). House dust endotoxin and wheeze in the first year of life. *American journal of respiratory and critical care medicine*, 163(2), 322–8. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11179100>

Park, B. S., Song, D. H., Kim, H. M., Choi, B.-S., Lee, H., & Lee, J.-O. (2009). The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature*, 458(7242), 1191–5. doi:10.1038/nature07830

Parker, J. M., Guo, D., & Hodges, R. S. (1986). New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry*, 25(19), 5425–32. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2430611>

Pearce, N., Ait-Khaled, N., Beasley, R., Mallol, J., Keil, U., Mitchell, Ed., Robertson, C., & I. P. T. S. G. (2007). Worldwide trends in the prevalence of asthma symptoms: phase III of the International Study of Asthma and Allergies in Childhood (ISAAC). *Thorax*, 62(9), 758–66. doi:10.1136/thx.2006.070169

Pellequer, J. L., Westhof, E., & Van Regenmortel, M. H. (1993). Correlation between the location of antigenic sites and the prediction of turns in proteins. *Immunology letters*, 36(1), 83–99. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7688347>

Piboonpocanun, S., Malainual, N., Jirapongsananuruk, O., Vichyanond, P., & Thomas, W. R. (2006). Genetic polymorphisms of major house dust mite allergens. *Clinical and Experimental Allergy*, 36(5), 510–516. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2222.2006.02464.x/full>

Pierce, B. G., Hourai, Y., & Weng, Z. (2011). Accelerating protein docking in ZDOCK using an advanced 3D convolution library. *PloS one*, 6(9), e24657. doi:10.1371/journal.pone.0024657

Pizzi, E., Cortese, R., & Tramontano, a. (1995). Mapping epitopes on protein surfaces. *Biopolymers*, 36(5), 675–80. doi:10.1002/bip.360360513

- Platts-Mills, T. A. E., & Woodfolk, J. A. (2011). Allergens and their role in the allergic immune response. *Immunological reviews*, 242(1), 51–68. doi:10.1111/j.1600-065X.2011.01021.x
- Pomés, A. (2002). Intrinsic properties of allergens and environmental exposure as determinants of allergenicity. *Allergy*, 57(8), 673–9S. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12121184>
- Pomés, A. (2010). Relevant B cell epitopes in allergic disease. *International archives of allergy and immunology*, 152(1), 1–11. doi:10.1159/000260078
- Ponomarenko, J. V., & Bourne, P. E. (2007). Antibody-protein interactions: benchmark datasets and prediction tools evaluation. *BMC structural biology*, 7, 64. doi:10.1186/1472-6807-7-64
- Ponomarenko, J., Bui, H-H., Li, W., Fusseder, N., Bourne, P. E., Sette, A., & Peters, B. (2008). ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC bioinformatics*, 9, 514. doi:10.1186/1471-2105-9-514
- Puerta, L., Caraballo, L., Fernández-Caldas, E., Avjioglu, A., Marsh, D. G., Lockey, R. F., & Dao, M. L. (1996). Nucleotide sequence analysis of a complementary DNA coding for a *Blomia tropicalis* allergen. *The Journal of allergy and clinical immunology*, 98(5 Pt 1), 932–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8939156>
- Punta, M., Coghill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E. L. L., Eddy, S. R., & Bateman, A., & Finn, R. D. (2012). The Pfam protein families database. *Nucleic acids research*, 40(Database issue), D290–301. doi:10.1093/nar/gkr1065

## R

- Radauer, C., Bublin, M., Wagner, S., Mari, A., & Breiteneder, H. (2008). Allergens are distributed into few protein families and possess a restricted number of biochemical functions. *The Journal of allergy and clinical immunology*, 121(4), 847–52.e7. doi:10.1016/j.jaci.2008.01.025
- Ramos, J. D., Cheong, N., Lee, B. W., Chua, K. Y., Nge, C., Wah, L. B., & Yan, C. K. (2001). cDNA cloning and expression of Blo t 11, the *Blomia tropicalis* allergen homologous to paramyosin. *International archives of allergy and immunology*, 126(4), 286–93. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11815735>
- Ramos, J. D. a, Cheong, N., Teo, a S. M., Kuo, I-C., Lee, B. W., & Chua, K. Y. (2004). Production of monoclonal antibodies for immunoaffinity purification and quantitation of Blo t 1 allergen in mite and dust extracts. *Clinical and experimental allergy : journal*

- of the British Society for Allergy and Clinical Immunology*, 34(4), 604–10. doi:10.1111/j.1365-2222.2004.1922.x
- Reese, G., Ballmer-Weber, B. K., Wangorsch, A., Randow, S., & Vieths, S. (2007). Allergenicity and antigenicity of wild-type and mutant, monomeric, and dimeric carrot major allergen Dau c 1: destruction of conformation, not oligomerization, is the roadmap to save allergen vaccines. *The Journal of allergy and clinical immunology*, 119(4), 944–51. doi:10.1016/j.jaci.2006.11.699
- Reginald, K., Gao, Y. F., Sew, Y. S., Shang, H. S., & Chew, F. T. (2004). Cross Comparison of the IgE Binding Profiles to Recombinant Allergens From *Suidasia medanensis*, *Blomia tropicalis* and *Dermatophagoides farinae* Using Sera From *Blomia*- and *Dermatophagoides*-Predominant Environments. *Journal of Allergy and Clinical Immunology*, 113(2), 228–229.
- Reginald, K. (2006). Groupe 2 allergens from dust mite: Epitope mapping and functional characterization of Der p 2, and identification of a paralogue of Der f 2. *Thèse de doctorat*.
- Rizzo, M. C. (2007). Endotoxin exposure and symptoms in asthmatic children . *Pediatric allergy and ...*, 9532251. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1399-3038.1997.tb00164.x/abstract>
- Rohl, C. A., Strauss, C. E. M., Chivian, D., & Baker, D. (2004). Modeling structurally variable regions in homologous proteins with rosetta. *Proteins*, 55(3), 656–77. doi:10.1002/prot.10629
- Rose, G. D., Geselowitz, a R., Lesser, G. J., Lee, R. H., & Zehfus, M. H. (1985). Hydrophobicity of amino acid residues in globular proteins. *Science (New York, N.Y.)*, 229(4716), 834–8. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/19334678>
- Rost, B. (1999). Twilight zone of protein sequence alignments. *Protein engineering*, 12(2), 85–94. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10195279>
- Roy, D., Ghosh, D., & Gupta-Bhattacharya, S. (2003). Homology modeling of allergenic cyclophilins: IgE-binding site and structural basis of cross-reactivity. *Biochemical and Biophysical Research Communications*, 307(2), 422–429. doi:10.1016/S0006-291X(03)01193-8
- Rubinstein, N. D., Mayrose, I., Halperin, D., Yekutieli, D., Gershoni, J. M., & Pupko, T. (2008). Computational characterization of B-cell epitopes. *Molecular immunology*, 45(12), 3477–89. doi:10.1016/j.molimm.2007.10.016

Rubinstein, N. D., Mayrose, I., Martz, E., & Pupko, T. (2009). Epitopia: a web-server for predicting B-cell epitopes. *BMC bioinformatics*, *10*, 287. doi:10.1186/1471-2105-10-287

## S

Saarne, T., Kaiser, L., Rasool, O., Huecas, S., van Hage-Hamsten, M., & Gafvelin, G. (2003). Cloning and characterisation of two IgE-binding proteins, homologous to tropomyosin and alpha-tubulin, from the mite *Lepidoglyphus destructor*. *International archives of allergy and immunology*, *130*(4), 258–65. doi:70212

Saha, S., & Raghava, G. (2004). BcePred: Prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. *Artificial Immune Systems*, *3239*, 197–204. Retrieved from <http://www.springerlink.com/index/ex360mj1wahflwmw.pdf>

Saha, S., & Raghava, G. P. S. (2006a). AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic acids research*, *34*(Web Server issue), W202–9. doi:10.1093/nar/gkl343

Saha, S., & Raghava, G. (2006b). Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *PROTEINS: Structure, Function, and Bioinformatics*, *65*(1), 31–33. doi:10.1002/prot.v65

Sali, A., & Blundell, T. (1993). Comparative Protein Modelling by Satisfaction of Spatial Restraints. *Protein Structure by Distance Analysis*, *234*, 779–815. Retrieved from [http://books.google.com/books?hl=en&lr=&id=LKxFL1ksNH8C&oi=fnd&pg=PA64&dq=Comparative+Protein+Modelling+by+Satisfaction+of+Spatial+Restraints&ots=gd7RSC5Vwf&sig=Z42X6\\_bvfobz\\_hnfLleJFBJoUd0](http://books.google.com/books?hl=en&lr=&id=LKxFL1ksNH8C&oi=fnd&pg=PA64&dq=Comparative+Protein+Modelling+by+Satisfaction+of+Spatial+Restraints&ots=gd7RSC5Vwf&sig=Z42X6_bvfobz_hnfLleJFBJoUd0)

Sánchez-Ramos, I., Alvarez-Alfageme, F., & Castañera, P. (2007). Effects of relative humidity on development, fecundity and survival of three storage mites. *Experimental & applied acarology*, *41*(1-2), 87–100. doi:10.1007/s10493-007-9052-7

Schein, C. H., Ivanciuc, O., & Braun, W. (2007). Bioinformatics approaches to classifying allergens and predicting cross-reactivity. *Immunology and allergy clinics of North America*, *27*(1), 1–27. doi:10.1016/j.iac.2006.11.005

Schlessinger, A., Ofran, Y., Yachdav, G., & Rost, B. (2006). Epitome: database of structure-inferred antigenic epitopes. *Nucleic acids research*, *34*(Database issue), D777–80. doi:10.1093/nar/gkj053

- Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., & Wolfson, H. J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic acids research*, 33(Web Server issue), W363–7. doi:10.1093/nar/gki481
- Schulz, O., Sutton, B. J., Beavil, R. L., Shi, J., Sewell, H. F., Gould, H. J., Laing, P., & Shakib, F. (1997). Cleavage of the low-affinity receptor for human IgE (CD23) by a mite cysteine protease: nature of the cleaved fragment in relation to the structure and function of CD23. *European journal of immunology*, 27(3), 584–8. doi:10.1002/eji.1830270303
- Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research*, 31(13), 3381–3385. doi:10.1093/nar/gkg520
- Schöll, I., Kalkura, N., Shedziankova, Y., Bergmann, A., Verdino, P., Knittelfelder, R., Kopp, T., Hantusch, B., Betzel, C., Dierks, K., Scheiner, O., Boltz-Nitulescu, G., Keller, W., Jensen-Jarolim, E. (2005). Dimerization of the major birch pollen allergen Bet v 1 is important for its in vivo IgE-cross-linking potential in mice. *Journal of immunology (Baltimore, Md. : 1950)*, 175(10), 6645–50. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16272319>
- Shakib, F., Ghaemmaghami, A. M., & Sewell, H. F. (2008). The molecular basis of allergenicity. *Trends in immunology*, 29(12), 633–42. doi:10.1016/j.it.2008.08.007
- Shek, L. P-C., Chong, A. R., Soh, S. E., Cheong, N., Teo, A. S. M., Yi, F. C., Giam, Y. C., Chua, K. Y., & Van Bever, H. P. (2010). Specific profiles of house dust mite sensitization in children with asthma and in children with eczema. *Pediatric allergy and immunology: official publication of the European Society of Pediatric Allergy and Immunology*, 21(4 Pt 2), e718–22. doi:10.1111/j.1399-3038.2010.01019.x
- Shen, H. D., Chua, K. Y., Lin, K. L., Hsieh, K. H., & Thomas, W. R. (1993). Molecular cloning of a house dust mite allergen with common antibody specificities with multiple components in mite extracts. *Clinical & Experimental Allergy*, 23(11), 934–40. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2222.1993.tb00278.x/abstract>
- Shen, M., & Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Science*, 15, 2507–2524. doi:10.1110/ps.062416606.Instead
- Shen, H-D., Tam, M. F., Huang, C-H., Chou, H., Tai, H-Y., Chen, Y-S., Sheu, S-Y., & Thomas, W. R. (2011). Homology modeling and monoclonal antibody binding of the Der f 7 dust mite allergen. *Immunology and cell biology*, 89(2), 225–30. doi:10.1038/icb.2010.77

- Simard, J. (2012). Collaboration haptique étroitement couplée pour la manipulation moléculaire interactive. *Thesis*. Retrieved from [http://hal.inria.fr/docs/00/68/80/36/PDF/VD2\\_Jean\\_SIMARD\\_12-03-2012.pdf](http://hal.inria.fr/docs/00/68/80/36/PDF/VD2_Jean_SIMARD_12-03-2012.pdf)
- Simpson, A., Green, R., Custovic, A., Woodcock, A., Arruda, L. K., & Chapman, M. D. (2003). Skin test reactivity to natural and recombinant *Blomia* and *Dermatophagoides* spp. allergens among mite allergic patients in the UK. *Allergy*, *58*(1), 53–6. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12580807>
- Siraganian, R. (2003). Mast cell signal transduction from the high-affinity IgE receptor. *Current Opinion in Immunology*, *15*(6), 639–646. doi:10.1016/j.coi.2003.09.010
- Smit, L. A. M., Heederik, D., Doekes, G., Blom, C., van Zweden, I., & Wouters, I. M. (2008). Exposure-response analysis of allergy and respiratory symptoms in endotoxin-exposed adults. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*, *31*(6), 1241–8. doi:10.1183/09031936.00090607
- Smith, W. a, Chua, K. Y., Kuo, M. C., Rogers, B. L., & Thomas, W. R. (1994). Cloning and sequencing of the *Dermatophagoides pteronyssinus* group III allergen, Der p III. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*, *24*(3), 220–8. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8911700>
- Smith, A. M., & Chapman, M. D. (1996). Reduction in IgE binding to allergen variants generated by site-directed mutagenesis: contribution of disulfide bonds to the antigenic structure of the major house dust mite allergen Der p 2. *Molecular immunology*, *33*(4-5), 399–405. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8676891>
- Smith, G., & Petrenko, V. (1997). Phage on display. *Chemistry & biology*, *97*(2), 391–410. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9444395>
- Smith, W., Mills, K., Hazell, L., Hart, B., & Thomas, W. (1999). Molecular analysis of the group 1 and 2 allergens from the house dust mite, *Euroglyphus maynei*. *International archives of allergy and immunology*, *118*(1), 15–22. doi:24026
- Smith, A. M., Benjamin, D. C., Hozic, N., Derewenda, U., Smith, W. a, Thomas, W. R., Gafvelin, G., van Hage-Hamsten, M., & Chapman, M. D. (2001). The molecular basis of antigenic cross-reactivity between the group 2 mite allergens. *The Journal of allergy and clinical immunology*, *107*(6), 977–84. doi:10.1067/mai.2001.115629
- Soto-Quiros, M., Avila, L., Platts-Mills, T. A. E., Hunt, J. F., Erdman, D. D., Carper, H., Murphy, D. D., Odio, S., James, H. R., Patrie, J. T., Hunt, W., O'Rourke, A. K., Davis, M. D., Steinke, J. W., Lu, X., Kennedy, J., & Heymann, P. W. (2012). High titers of IgE antibody to dust mite allergen and risk for wheezing among asthmatic children infected

- with rhinovirus. *The Journal of allergy and clinical immunology*, 129(6), 1499–1505.e5. doi:10.1016/j.jaci.2012.03.040
- Spieksma, F. T. (1997). Domestic mites from an acarologic perspective. *Allergy*, 52(4), 360–8. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9188914>
- Stewart, G. A., Bird, C. H., Krska, K. D., Colloff, M. J., & Thompson, P. J. (1992). A comparative study of allergenic and potentially allergenic enzymes from *Dermatophagoides pteronyssinus*, *D. farinae* and *Euroglyphus maynei*. *Experimental & applied acarology*, 16(1-2), 165–80. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1283568>
- Strachan, D., Sibbald, B., Weiland, S., Aït-Khaled, N., Anabwani, G., Anderson, H. R., Asher, M. I., Beasley, R., Björkstén, B., Burr, M., Clayton, T., Crane, J., Ellwood, P., Keil, U., Lai, C., Mallol, J., Martinez, F., & Mi, H. (1997). Worldwide variations in prevalence of symptoms of allergic rhinoconjunctivitis in children: the International Study of Asthma and Allergies in Childhood (ISAAC). *Pediatric allergy and immunology: official publication of the European Society of Pediatric Allergy and Immunology*, 8, 161–176. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9553981>
- Sun, J., Wu, D., Xu, T., Wang, X., Xu, X., Tao, L., Li, Y. X., & Cao, Z. W. (2009). SEPPA: a computational server for spatial epitope prediction of protein antigens. *Nucleic acids research*, 37(Web Server issue), W612–6. doi:10.1093/nar/gkp417
- Sun, P., Chen, W., Huang, Y., Wang, H., Ma, Z., & Lv, Y. (2011). Epitope prediction based on random peptide library screening: benchmark dataset and prediction tools evaluation. *Molecules*, 4971–4993. doi:10.3390/molecules16064971
- Surdu, S., Montoya, L. D., Tarbell, A., & Carpenter, D. O. (2006). Childhood asthma and indoor allergens in Native Americans in New York. *Environmental health: a global access science source*, 5, 22. doi:10.1186/1476-069X-5-22
- Suzuki, M., Tanaka, Y., Korematsu, S., Mikami, B., & Minato, N. (2006). Crystal structure and some properties of a major house dust mite allergen, Derf 2. *Biochemical and biophysical research communications*, 339(2), 679–86. doi:10.1016/j.bbrc.2005.11.065
- Sweredoski, M. J., & Baldi, P. (2008). PEPITO: improved discontinuous B-cell epitope prediction using multiple distance thresholds and half sphere exposure. *Bioinformatics (Oxford, England)*, 24(12), 1459–60. doi:10.1093/bioinformatics/btn199
- Szalai, K., Fuhrmann, J., Pavkov, T., Scheidl, M., Wallmann, J., Brämshwig, K. H., Vrtala, S., Scheiner, O., Keller, W., Saint-Remy, J-M., Neumann, D., Pali-Schöll, I., & Jensen-Jarolim, E. (2008). Mimotopes identify conformational B-cell epitopes on the two major

house dust mite allergens Der p 1 and Der p 2. *Molecular immunology*, 45(5), 1308–17. doi:10.1016/j.molimm.2007.09.012

## T

- Takai, T., Yuuki, T., Okumura, Y., Mori, A., & Okudaira, H. (1997). Determination of the N- and C-terminal sequences required to bind human IgE of the major house dust mite allergen Der f 2 and epitope mapping for monoclonal antibodies. *Molecular immunology*, 34(3), 255–261.
- Takai, Toshiro, Kato, T., Sakata, Y., Yasueda, H., Izuhara, K., Okumura, K., & Ogawa, H. (2005). Recombinant Der p 1 and Der f 1 exhibit cysteine protease activity but no serine protease activity. *Biochemical and biophysical research communications*, 328(4), 944–52. doi:10.1016/j.bbrc.2005.01.051
- Tan, K. W., Kumar, S., Chew, F. T., & Mok, Y. K. (2011). Cloning, expression, purification, crystallization and preliminary X-ray diffraction studies of a major group 7 allergen, Der f 7, from the dust mite *Dermatophagoides farinae*. *Acta crystallographica. Section F, Structural biology and crystallization communications*, 67(Pt 12), 1612–5. doi:10.1107/S174430911103836X
- Tan, K. W., Ong, T. C., Gao, Y. F., Tiong, Y. S., Wong, K. N., Chew, F. T., & Mok, Y. K. (2012). NMR Structure and IgE Epitopes of Blo t 21, a Major Dust Mite Allergen from *Blomia tropicalis*. *The Journal of biological chemistry*, 287(41), 34776–85. doi:10.1074/jbc.M112.348730
- Tang, J. C., Wong, S. F., Mak, J. W., & Ho, T. M. (2011). Antigenic profile of *Blomia tropicalis*, *Aleuroglyphus ovatus* and *Glycycometus malaysiensis*. *Tropical biomedicine*, 28(2), 223–36. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/22041741>
- Tategaki, A., Kawamoto, S., Ski, T., Jyo, T., Suzuki, O., Shigeta, S., & Ono, K. (2000). Newly described house dust mite allergens. *Journal of Allergy and Clinical Immunology*, 1, 74–76. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Newly+described+house+dust+mite+allergens#0>
- The Uniprot Consortium. (2012). Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic acids research*, 40(Database issue), D71–5. doi:10.1093/nar/gkr981
- Thomas, M. (2006). Allergic rhinitis: evidence for impact on asthma. *BMC pulmonary medicine*, 6 Suppl 1, S4. doi:10.1186/1471-2466-6-S1-S4

- Thomas, W. R., Smith, W.-A., Hales, B. J., Mills, K. L., & O'Brien, R. M. (2002). Characterization and immunobiology of house dust mite allergens. *International archives of allergy and immunology*, *129*(1), 1–18. doi:65179
- Thomas, W. R., Heinrich, T. K., Smith, W. A., & Hales, B. J. (2007). Pyroglyphid House Dust Mite Allergens. *Protein and peptid Letters*, *14*(10), 943–953. Retrieved from <http://www.ingentaconnect.com/content/ben/ppl/2007/00000014/00000010/art00003>
- Thomas, W. R., Hales, B. J., & Smith, W.-A. (2010). House dust mite allergens in asthma and allergy. *Trends in molecular medicine*, *16*(7), 321–8. doi:10.1016/j.molmed.2010.04.008
- Thomas, W. R. (2011). The advent of recombinant allergens and allergen cloning. *The Journal of allergy and clinical immunology*, *127*(4), 855–9. doi:10.1016/j.jaci.2010.12.1084
- Thomas, W. R. (2012). House dust allergy and immunotherapy. *Human vaccines & immunotherapeutics*, *8*(10), 1–10. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/22894952>
- Thorn, J., & Rylander, R. (1998). Inflammatory response after inhalation of bacterial endotoxin assessed by the induced sputum technique. *Thorax*, *53*(12), 1047–52. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1745135&tool=pmcentrez&rendertype=abstract>
- Tina, K. G., Bhadra, R., & Srinivasan, N. (2007). PIC: Protein Interactions Calculator. *Nucleic acids research*, *35*(Web Server issue), W473–6. doi:10.1093/nar/gkm423
- Tomee, J. F., van Weissenbruch, R., de Monchy, J. G., & Kauffman, H. F. (1998). Interactions between inhalant allergen extracts and airway epithelial cells: effect on cytokine production and cell detachment. *The Journal of allergy and clinical immunology*, *102*(1), 75–85. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9679850>
- Tovchigrechko, A., & Vakser, I. A. (2006). GRAMM-X public web server for protein-protein docking. *Nucleic acids research*, *34*(Web Server issue), W310–4. doi:10.1093/nar/gkl206
- Trompette, A., Divanovic, S., Visintin, A., Blanchard, C., Rashmi, S., Madan, R., Thorne, P. S., Wills-karp, M., Gioannini, T. L., Weiss, P., & Karp, C. L. (2009). Allergenicity resulting from functional mimicry of a Toll-like receptor complex protein. *Nature*, *457*(7229), 585–588. doi:10.1038/nature07548.Allergenicity

- Trudinger, M., Chua, K. Y., & Thomas, W. R. (1991). cDNA encoding the major mite allergen Der f II. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*, 21(1), 33–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2021876>
- Tsai, L. C., Chao, P. L., Shen, H. D., Tang, R. B., Chang, T. C., Chang, Z. N., Hung, M. W., Lee, B. L., & Chua, K. Y. (1998). Isolation and characterization of a novel 98-kd *Dermatophagoides farinae* mite allergen. *The Journal of allergy and clinical immunology*, 102(2), 295–303. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9723675>
- Tsai, J., Yi, F., Chua, K., & Liu, Y. (2003). Identification of the major allergenic components in *Blomia tropicalis* and the relevance of the specific IgE in asthmatic patients. *Annals of allergy, asthma & immunology*, 2, 11–12. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1081120610615189>
- Tsai, J. J., Liu, S. H., Yin, S. C., Yang, C. N., Hsu, H. S., Chen, W. B., Liao, E. C., Lee, W. J., Pan, H. C., & Sheu, and M. L. (2011). Mite allergen Der-p2 triggers human B lymphocyte activation and Toll-like receptor-4 induction. *PloS one*, 6(9), e23249. doi:10.1371/journal.pone.0023249

## V

- Valenta, R. (2002). Recombinant allergen-based concepts for diagnosis and therapy of type I allergy. *Allergy, 57 Suppl 7*, 66–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12173274>
- Valenta, R., Linhart, B., Swoboda, I., & Niederberger, V. (2011). Recombinant allergens for allergen-specific immunotherapy: 10 years anniversary of immunotherapy with recombinant allergens. *Allergy*, 66(6), 775–83. doi:10.1111/j.1398-9995.2011.02565.x
- Valerio, C. R., Murray, P., Arlian, L. G., & Slater, J. E. (2005). Bacterial 16S ribosomal DNA in house dust mite cultures. *The Journal of allergy and clinical immunology*, 116(6), 1296–300. doi:10.1016/j.jaci.2005.09.046
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. C. (2005). GROMACS: fast, flexible, and free. *Journal of computational chemistry*, 26(16), 1701–18. doi:10.1002/jcc.20291
- Van Regenmortel, M. H. (2001). Antigenicity and immunogenicity of synthetic peptides. *Biologicals : journal of the International Association of Biological Standardization*, 29(3–4), 209–13. doi:10.1006/biol.2001.0308

- Vihinen, M., Torkkila, E., & Riikonen, P. (1994). Accuracy of protein flexibility predictions. *Proteins*, *19*(2), 141–9. doi:10.1002/prot.340190207
- Vita, R., Zarebski, L., Greenbaum, J. A., Emami, H., Hoof, I., Salimi, N., Damle, R., Sette, A., & Peters, B. (2010). The immune epitope database 2.0. *Nucleic acids research*, *38*(Database issue), D854–62. doi:10.1093/nar/gkp1004
- Vrtala, S., Hirtenlehner, K., Susani, M., Akdis, M., Kussebi, F., Akdis, C. a, Blaser, K., Hufnagl, P., Binder, B. R., Politou, A., Pastore, A., Vangelista, L., Sperr, W. R., Semper, H., Valent, P., Ebner, C., Kraft, D., & Valenta, R. (2001). Genetic engineering of a hypoallergenic trimer of the major birch pollen allergen Bet v 1. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, *15*(11), 2045–7. doi:10.1096/fj.00-0767fje

## W

- Walgraffe, D., Mattéotti, C., El Bakkoury, M., Garcia, L., Marchand, C., Bullens, D., Vandenbranden, M., & Jacquet, A. (2009). A hypoallergenic variant of Der p 1 as a candidate for mite allergy vaccines. *The Journal of allergy and clinical immunology*, *123*(5), 1150–6. doi:10.1016/j.jaci.2008.11.038
- Wan, H., Winton, H. L., Soeller, C., Gruenert, D. C., Thompson, P. J., Cannell, M. B., Stewart, G. A., Garrod, D. R., & Robinson, C. (2000). Quantitative structural and biochemical analyses of tight junction dynamics following exposure of epithelial cells to house dust mite allergen Der p 1. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*, *30*(5), 685–98. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10792361>
- Wang, Y.-H., & Liu, Y.-J. (2008). The IL-17 cytokine family and their role in allergic inflammation. *Current opinion in immunology*, *20*(6), 697–702. doi:10.1016/j.coi.2008.09.004
- Weber, E., Hunter, S., Stedman, K., Dreitz, S., Olivry, T., Hillier, A., & McCall, C. (2003). Identification, characterization, and cloning of a complementary DNA encoding a 60-kd house dust mite allergen (Der f 18) for human beings and dogs. *Journal of Allergy and Clinical Immunology*, *112*(1), 79–86. doi:10.1067/mai.2003.1602
- Weghofer, M., Thomas, W. R., Kronqvist, M., Mari, A., Purohit, A., Pauli, G., Horak, F., Grönlund, H., van Hage, M., Valenta, R., & Vrtala, S. (2008a). Variability of IgE reactivity profiles among European mite allergic patients. *European journal of clinical investigation*, *38*(12), 959–65. doi:10.1111/j.1365-2362.2008.02048.x
- Weghofer, M., Grote, M., Dall'Antonia, Y., Fernández-Caldas, E., Krauth, M. T., van Hage, M., Horak, F., Thomas, W. R., Valent, P., Keller, W., Valenta, R., & Vrtala, S. (2008b).

Characterization of folded recombinant Der p 5, a potential diagnostic marker allergen for house dust mite allergy. *International archives of allergy and immunology*, 147(2), 101–9. doi:10.1159/000135696

Weghofer, M., Dall'Antonia, Y., Grote, M., Stöcklinger, A., Kneidinger, M., Balic, N., Krauth, M-T., Fernández-Caldas, E., Thomas, W. R., van Hage, M., Vieths, S., Spitzauer, S., Horak, F., Svergun, D. I., Konarev, P. V., Valent, P., Thalhamer, J., Keller, S. (2008c). Characterization of Der p 21, a new important allergen derived from the gut of house dust mites. *Allergy*, 63(6), 758–67. doi:10.1111/j.1398-9995.2008.01647.x

Weiland, S. K., Björkstén, B., Brunekreef, B., Cookson, W. O. C., von Mutius, E., & Strachan, D. P. (2004). Phase II of the International Study of Asthma and Allergies in Childhood (ISAAC II): rationale and methods. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*, 24(3), 406–12. doi:10.1183/09031936.04.00090303

Wills-Karp, M. (2010). Allergen-specific pattern recognition receptor pathways. *Current opinion in immunology*, 22(6), 777–82. doi:10.1016/j.coi.2010.10.011

Witteman, A. M., Akkerdaas, J. H., van Leeuwen, J., van der Zee, J. S., & Aalberse, R. C. (1994). Identification of a Cross-Reactive Allergen (Presumably Tropomyosin) in Shrimp, Mite and Insects. *International Archives of Allergy and Immunology*, 105(1), 56–61. doi:10.1159/000236803

Wong, S. F., Chong, A. L., Mak, J. W., Tan, J., Ling, S. J., & Ho, T. M. (2011). Molecular identification of house dust mites and storage mites. *Experimental & applied acarology*, 55(2), 123–33. doi:10.1007/s10493-011-9460-6

## X

Xu, X., Sun, J., Liu, Q., Wang, X., Xu, T., Zhu, R., Wu, D., & Cao, Z. (2010). Evaluation of spatial epitope computational tools based on experimentally-confirmed dataset for protein antigens. *Chinese Science Bulletin*, 55(20), 2169–2174. doi:10.1007/s11434-010-3199-z

## Y

Yadav, A., Elder, B. L., Morgan, M. S., Vyszynski-Moher, D. L., & Arlian, L. G. (2006). Prevalence of serum IgE to storage mites in a southwestern Ohio population. *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology*, 96(2), 356–62. doi:10.1016/S1081-1206(10)61248-3

Yi, F. C., Cheong, N., Shek, P. C., Wang D. Y., Chua K.Y., Lee, B. (2002). Identification of shared and unique immunoglobulin E epitopes of the highly conserved tropomyosins in *Blomia tropicalis* and *Dermatophagoides pteronyssinus*. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*, 32(8), 1203–10. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12190660>

Yi, F. C., Chen, J. Y., Chee, K. K., Chua, K. Y., & Lee, B. W. (2009). Dust mite infestation of flour samples. *Allergy*, 64(12), 1788–9. doi:10.1111/j.1398-9995.2009.02116.x

## Z

Zakzuk, J., Jiménez, S., Cheong, N., Puerta, L., Lee, B. W., Chua, K. Y., & Caraballo, L. (2009). Immunological characterization of a Blo t 12 isoallergen: identification of immunoglobulin E epitopes. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*, 39(4), 608–16. doi:10.1111/j.1365-2222.2008.03193.x

Zhang, C., Gjesing, B., Lai, X., Li, J., Spangfort, M. D., & Zhong, N. (2011). Indoor allergen levels in Guangzhou city, southern China. *Allergy*, 66(2), 186–91. doi:10.1111/j.1398-9995.2010.02465.x

Zhang, L., Huang, Y., Zou, Z., He, Y., Chen, X., & Tao, A. (2012a). SORTALLER: predicting allergens using substantially optimized algorithm on allergen family featured peptides. *Bioinformatics* (Oxford, England), 28(16), 2178–9. doi:10.1093/bioinformatics/bts326

Zhang, W., Niu, Y., Xiong, Y., Zhao, M., Yu, R., & Liu, J. (2012b). Computational prediction of conformational B-cell epitopes from antigen primary structures by ensemble learning. *PLoS one*, 7(8), e43575. doi:10.1371/journal.pone.0043575

Zhao, J., Lloyd, C., & Noble, A. (2012). Th17 responses in chronic allergic airway inflammation abrogate regulatory T-cell-mediated tolerance and contribute to airway remodeling. *Mucosal Immunology*. doi:10.1038/mi.2012.76

Zhong, Z., Zheng, C., & Wang, F. (2011). Identification of the IgE-binding epitopes in main dust mite allergen Der p 1. *Nan fang yi ke da xue xue bao = Journal of Southern Medical University*, 31(7), 1183–6. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/21764690>

Zhou, H. X., & Qin, S. (2007). Interaction-site prediction for protein complexes: a critical assessment. *Bioinformatics* (Oxford, England), 23(17), 2203–9. doi:10.1093/bioinformatics/btm323

- Zock, J. P., Brunekreef, B., Hazebroek-Kampschreur, A. A. J. M., & Roosjen, C. W. (1994). House dust mite allergen in bedroom floor dust and respiratory health of children with asthmatic symptoms. *European Respiratory Journal*, 7(7), 1254–1259. doi:10.1183/09031936.94.07071254

# *Annexes*

## Annexe I: Scripts pour la modélisation comparative et la minimisation d'énergie des modèles

Fichiers utilisés pour la modélisation de la structure de Der f 5 (exemple) sur la base de la structure expérimentale de Der p 5 par le programme MODELLER:

### *Le fichier d'entrée 'Derf5.ali'*

```
>P1;Derf5
sequence:Derf5:  ::  :::: 0.00: 0.00
CVSGEPKKHDYQNEFDLLMQRIHEQMRKGEEALLHLQHQINTFEENPTKEMKEQILGEMDTIHALIDGVRGV
LNRLMKRTDLDFIFERYNVEIALKSNEILERDLKKEEQRVKKIEV*
```

### *Le fichier de commande 'align2d.py'*

```
from modeller import *

env = environ()
aln = alignment(env)
mdl = model(env, file='3MQ1', model_segment=('FIRST:A','LAST:A'))
aln.append_model(mdl, align_codes='3MQ1A', atom_files='3MQ1.pdb')
aln.append(file='Derf5.ali', align_codes='Derf5')
aln.align2d()
aln.write(file='Derf5-Derp5.ali', alignment_format='PIR')
aln.write(file='Derf5-Derp5.pap', alignment_format='PAP')
```

### *Le fichier de sortie 'Derf5-Derp5.ali'*

```
>P1;3MQ1A
structureX:3MQ1.pdb: 32 :A:+101 :A:MOL_ID 1; MOLECULE MITE ALLERGEN DER P 5; CHAIN A, B,
C, D, E, F; FRAGMENT UNP RESIDUES 34-132; SYNONYM ALLERGEN DER P V, IGE-BINDING
ALLERGEN; ENGINEERED YES:MOL_ID 1; ORGANISM_SCIENTIFIC DERMATOPHAGOIDES
PTERONYSSINUS; ORGANISM_COMMON EUROPEAN HOUSE DUST MITE; ORGANISM_TAXID 6956;
GENE DERP5; EXPRESSION_SYSTEM ESCHERICHIA COLI; EXPRESSION_SYSTEM_TAXID 469008;
EXPRESSION_SYSTEM_STRAIN BL21(DE3)RIL; EXPRESSION_SYSTEM_VECTOR_TYPE PLASMID;
EXPRESSION_SYSTEM_PLASMID PGEX4T3(TEV): 2.00:-1.00
-----NSLMERIHEQIKKGELALFYLQEQINHFEEKPTKEMKDKIVAEMDTIHAMIDGVRGVLD
RLMQRKDLDFEQYNLEMAKKSGDILERDLKKEEARVKKIEV*

>P1;Derf5
sequence:Derf5:  ::  :::: 0.00: 0.00
CVSGEPKKHDYQNEFDLLMQRIHEQMRKGEEALLHLQHQINTFEENPTKEMKEQILGEMDTIHALIDGVRGV
LNRLMKRTDLDFIFERYNVEIALKSNEILERDLKKEEQRVKKIEV*
```

### *Le fichier de commande 'model-single.py'*

```
from modeller import *
from modeller.automodel import *

env = environ()
a = automodel(env, alnfile='Derf5-Derp5.ali',
              knowns='3MQ1A', sequence='Derf5')
```

Fichiers utilisés pour la minimisation d'énergies par Gromos des modèles de Der f 5 (exemple) générés par le programme MODELLER:

***Le fichier de commande 'batch\_pdb2gmx.csh'***

```

pdb2gmx -f $1.pdb -o $1.gro -p $1.top << EOF
2
EOF
grompp -f em_vide.mdp -c $1.gro -p $1.top -o $1_dp5.tpr
mdrun -s $1_dp5.tpr -o $1_dp5.trr -c $1_dp5.gro -g sd.log -e $1_dp5.edr
g_energy -f $1_dp5.edr -o $1_dp5.xvg << EOF
Potential
EOF
editconf -f $1_dp5.gro -o $1_dp5.pdb

```

***Le fichier d'entrée 'em\_vide.mdp'***

```

; LINES STARTING WITH ';' ARE COMMENTS
title           = Minimization of BPTI ()           ; Title of run

; The following lines tell the program the standard locations where to find certain files
cpp             = /usr/bin/cpp ; Preprocessor
define         =

; Parameters describing what to do, when to stop and what to save
integrator      = steep           ; Algorithm (steep = steepest descent minimization)
emtol          = 500.0           ; Stop minimization when the maximum force < 1800.0 kJ/mol
emstep         = 0.01
nsteps         = 200             ; Maximum number of (minimization) steps to perform
nstenergy      = 10             ; Write energies to disk every nstenergy steps
nstxtcout      = 10             ; Write coordinates to disk every nstxtcout steps
xtc_grps= Protein           ; Which coordinate group(s) to write to disk
energygrps     = Protein        ; Which energy group(s) to write to disk

; Parameters describing how to find the neighbors of each atom and how to calculate the interactions
pbc            = no
nstlist        = 5              ; Frequency to update the neighbor list and long range forces
ns_type        = simple         ; Method to determine neighbor list (simple, grid)
rlist          = 1.0            ; Cut-off for making neighbor list (short range forces)
coulombtype    = cut-off        ; Treatment of long range electrostatic interactions
rcoulomb       = 1.0            ; long range electrostatic cut-off
rvdw           = 1.0            ; long range Van der Waals cut-off
constraints    = none           ; Bond types to replace by constraints
Tcoupl         = no
Pcoupl         = no
gen_vel        = no

```

## Annexe II

**Tableau A1:** Score des modèles de monomères calculées en utilisant 2JMH comme patron (Blo t 5) et présentant une moyenne de z-score inférieure à -0,5. La plus basse énergie est en gras.

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
derf5_monomer <sub>1</sub> [2JMH]	-179.0	<b>-9805.5</b>	-1.06	-1.75	<b>-1.40</b>
derf5_monomer <sub>2</sub> [2JMH]	-177.0	-8726.2	-1.01	-0.99	-1.00
derf5_monomer <sub>3</sub> [2JMH]	-187.0	-8405.2	-1.22	-0.76	-0.99
derf5_monomer <sub>4</sub> [2JMH]	-156.0	-9133.6	-0.59	-1.27	-0.93
derf5_monomer <sub>5</sub> [2JMH]	-184.0	-8051.5	-1.16	-0.51	-0.83
derf5_monomer <sub>6</sub> [2JMH]	<b>-190.0</b>	-7867.4	-1.28	-0.38	-0.83
derf5_monomer <sub>7</sub> [2JMH]	-148.0	-8927.2	-0.43	-1.13	-0.78
derf5_monomer <sub>8</sub> [2JMH]	-181.0	-7718.4	-1.10	-0.27	-0.69
derf5_monomer <sub>9</sub> [2JMH]	-166.0	-7831.8	-0.79	-0.35	-0.57
derf5_monomer <sub>10</sub> [2JMH]	-165.0	-7812.4	-0.77	-0.34	-0.56
Moyenne	-173.3	-8427.9			
écart-type	13.9	701.4			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
suim5_monomer <sub>1</sub> [2JMH]	-99.0	<b>-9078.4</b>	-1.21	-2.50	<b>-1.86</b>
suim5_monomer <sub>2</sub> [2JMH]	-75.0	-8262.3	-0.74	-1.66	-1.20
suim5_monomer <sub>3</sub> [2JMH]	-80.0	-8031.8	-0.84	-1.42	-1.13
suim5_monomer <sub>4</sub> [2JMH]	-66.0	-7921.9	-0.56	-1.31	-0.93
suim5_monomer <sub>5</sub> [2JMH]	-80.0	-7578.8	-0.84	-0.95	-0.89
suim5_monomer <sub>6</sub> [2JMH]	-87.0	-7211.8	-0.98	-0.57	-0.77
suim5_monomer <sub>7</sub> [2JMH]	-45.0	-7853.4	-0.15	-1.23	-0.69
suim5_monomer <sub>8</sub> [2JMH]	<b>-111.0</b>	-6551.6	-1.45	0.11	-0.67
suim5_monomer <sub>9</sub> [2JMH]	-75.0	-7197.8	-0.74	-0.56	-0.65
suim5_monomer <sub>10</sub> [2JMH]	-75.0	-6932.5	-0.74	-0.28	-0.51
Moyenne	-79.3	-7662.0			
écart-type	19.8	730.1			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
lepd5_monomer <sub>1</sub> [2JMH]	-231.0	<b>-8320.8</b>	-1.81	-1.67	<b>-1.74</b>
lepd5_monomer <sub>2</sub> [2JMH]	<b>-232.0</b>	-7711.8	-1.83	-1.10	-1.46
lepd5_monomer <sub>3</sub> [2JMH]	-211.0	-7832.9	-1.45	-1.21	-1.33
lepd5_monomer <sub>4</sub> [2JMH]	-166.0	-8254.4	-0.64	-1.60	-1.12
lepd5_monomer <sub>5</sub> [2JMH]	-229.0	-6897.4	-1.78	-0.33	-1.06
lepd5_monomer <sub>6</sub> [2JMH]	-161.0	-7682.9	-0.54	-1.07	-0.81
lepd5_monomer <sub>7</sub> [2JMH]	-145.0	-7846.3	-0.25	-1.22	-0.74
lepd5_monomer <sub>8</sub> [2JMH]	-147.0	-7741.2	-0.29	-1.12	-0.71
lepd5_monomer <sub>9</sub> [2JMH]	-156.0	-7445.6	-0.45	-0.85	-0.65
Moyenne	-186.4	-7748.1			
écart-type	38.3	421.6			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
derp21_monomer <sub>1</sub> [2JMH]	<b>-109.0</b>	-8498.8	-1.79	-1.51	<b>-1.65</b>
derp21_monomer <sub>2</sub> [2JMH]	-84.0	-8672.4	-1.36	-1.64	-1.50
derp21_monomer <sub>3</sub> [2JMH]	-100.0	-7587.4	-1.64	-0.80	-1.22
derp21_monomer <sub>4</sub> [2JMH]	-106.0	-7077.4	-1.74	-0.40	-1.07
derp21_monomer <sub>5</sub> [2JMH]	-15.0	<b>-8963.9</b>	-0.19	-1.87	-1.03
derp21_monomer <sub>6</sub> [2JMH]	-78.0	-7367.3	-1.26	-0.63	-0.95
derp21_monomer <sub>7</sub> [2JMH]	-15.0	-8209.1	-0.19	-1.28	-0.74
derp21_monomer <sub>8</sub> [2JMH]	-66.0	-6967.3	-1.06	-0.32	-0.69
derp21_monomer <sub>9</sub> [2JMH]	-77.0	-6658.9	-1.25	-0.08	-0.66
Moyenne	-72.2	-7778.1			
écart-type	35.4	830.5			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
derf21_monomer <sub>1</sub> [2JMH]	<b>-173.0</b>	<b>-7297.1</b>	-2.55	-2.17	<b>-2.36</b>
derf21_monomer <sub>2</sub> [2JMH]	-108.0	-6914.9	-1.34	-1.89	-1.62
derf21_monomer <sub>3</sub> [2JMH]	-104.0	-5659.1	-1.27	-0.99	-1.13
derf21_monomer <sub>4</sub> [2JMH]	-115.0	-5063.0	-1.47	-0.56	-1.02
derf21_monomer <sub>5</sub> [2JMH]	-74.0	-6011.9	-0.71	-1.24	-0.98
derf21_monomer <sub>6</sub> [2JMH]	-95.0	-4901.1	-1.10	-0.45	-0.77
derf21_monomer <sub>7</sub> [2JMH]	-81.0	-5175.9	-0.84	-0.64	-0.74
derf21_monomer <sub>8</sub> [2JMH]	-58.0	-5304.3	-0.41	-0.74	-0.57
Moyenne	-101.0	-5790.9			
écart-type	34.7	889.2			

**Tableau A2:** Score des modèles de monomères calculées en utilisant 2LM9 comme patron (Blo t 21) et présentant une moyenne de z-score inférieure à -0,5. La plus basse énergie est en gras.

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
derf5_monomer <sub>1</sub> [2LM9]	-142,0	-6837,9	-1,0	-0,2	-0,6
derf5_monomer <sub>2</sub> [2LM9]	-129,0	-7243,5	-0,5	-0,6	-0,6
derf5_monomer <sub>3</sub> [2LM9]	-129,0	-7190,8	-0,5	-0,6	-0,5
derf5_monomer <sub>4</sub> [2LM9]	-143,0	-8595,6	-1,0	-1,9	-1,5
derf5_monomer <sub>5</sub> [2LM9]	-139,0	-6776,7	-0,9	-0,2	-0,5
derf5_monomer <sub>6</sub> [2LM9]	-149,0	-6630,3	-1,3	0,0	-0,6
derf5_monomer <sub>7</sub> [2LM9]	<b>-160,0</b>	-6538,8	-1,7	0,1	-0,8
derf5_monomer <sub>8</sub> [2LM9]	-131,0	-7179,0	-0,6	-0,6	-0,6
derf5_monomer <sub>9</sub> [2LM9]	-124,0	<b>-9360,6</b>	-0,3	-2,7	<b>-1,5</b>
Moyenne	-138,4	-7372,6			
écart-type	11,5	963,2			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
suim5_monomer <sub>1</sub> [2LM9]	-145,0	<b>-7549,0</b>	-0,7	-1,7	-1,2
suim5_monomer <sub>2</sub> [2LM9]	<b>-204,0</b>	-6726,4	-2,9	-0,3	-1,6
suim5_monomer <sub>3</sub> [2LM9]	-142,0	-6944,3	-0,6	-0,7	-0,6
suim5_monomer <sub>4</sub> [2LM9]	-149,0	-6921,4	-0,8	-0,6	-0,7
suim5_monomer <sub>5</sub> [2LM9]	-186,0	-7147,1	-2,2	-1,0	<b>-1,6</b>
suim5_monomer <sub>6</sub> [2LM9]	-113,0	-7535,3	0,5	-1,6	-0,6
suim5_monomer <sub>7</sub> [2LM9]	-158,0	-7311,7	-1,2	-1,3	-1,2
suim5_monomer <sub>8</sub> [2LM9]	-145,0	-6846,0	-0,7	-0,5	-0,6
Moyenne	-155,3	-7122,7			
écart-type	28,1	314,9			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
lepd5_monomer <sub>1</sub> [2LM9]	-186,0	-6083,5	-0,9	-1,2	-1,1
lepd5_monomer <sub>2</sub> [2LM9]	-172,0	-5902,7	-0,6	-0,9	-0,8
lepd5_monomer <sub>3</sub> [2LM9]	-185,0	-5865,4	-0,9	-0,9	-0,9
lepd5_monomer <sub>4</sub> [2LM9]	-175,0	-6101,1	-0,7	-1,2	-1,0
lepd5_monomer <sub>5</sub> [2LM9]	<b>-192,0</b>	-5239,8	-1,1	0,1	-0,5
lepd5_monomer <sub>6</sub> [2LM9]	-186,0	-5366,4	-0,9	-0,1	-0,5
lepd5_monomer <sub>7</sub> [2LM9]	-147,0	<b>-6939,3</b>	0,0	-2,5	<b>-1,3</b>
Moyenne	-177,6	-5928,3			
écart-type	15,2	558,3			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
derp21_monomer <sub>1</sub> [2LM9]	-152,0	-6397,9	-1,0	-0,8	-0,9
derp21_monomer <sub>2</sub> [2LM9]	-150,0	-7023,7	-0,9	-1,3	-1,1
derp21_monomer <sub>3</sub> [2LM9]	-141,0	-6109,0	-0,7	-0,5	-0,6
derp21_monomer <sub>4</sub> [2LM9]	-138,0	-6062,0	-0,6	-0,5	-0,5
derp21_monomer <sub>5</sub> [2LM9]	-109,0	-6984,2	0,2	-1,3	-0,5

derp21_monomer <sub>6</sub> [2LM9]	-149,0	-5715,5	-0,9	-0,2	-0,5
derp21_monomer <sub>7</sub> [2LM9]	<b>-182,0</b>	-6433,7	-1,8	-0,8	-1,3
derp21_monomer <sub>8</sub> [2LM9]	-148,0	-6039,5	-0,9	-0,4	-0,7
derp21_monomer <sub>9</sub> [2LM9]	-130,0	<b>-8819,8</b>	-0,4	-3,0	<b>-1,7</b>
Moyenne	-144,3	-6620,6			
écart-type	19,5	931,8			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
derf21_monomer <sub>1</sub> [2LM9]	-165,0	-5774,0	-1,1	-0,9	-1,0
derf21_monomer <sub>2</sub> [2LM9]	-148,0	-5859,1	-0,6	-0,9	-0,8
derf21_monomer <sub>3</sub> [2LM9]	-163,0	<b>-7237,4</b>	-1,0	-2,2	-1,6
derf21_monomer <sub>4</sub> [2LM9]	-179,0	-4460,1	-1,4	0,3	-0,6
derf21_monomer <sub>5</sub> [2LM9]	-176,0	-5755,9	-1,4	-0,9	-1,1
derf21_monomer <sub>6</sub> [2LM9]	-170,0	-5472,6	-1,2	-0,6	-0,9
derf21_monomer <sub>7</sub> [2LM9]	-138,0	-6098,9	-0,4	-1,2	-0,8
derf21_monomer <sub>8</sub> [2LM9]	<b>-206,0</b>	-6679,3	-2,2	-1,7	<b>-1,9</b>
Moyenne	-168,1	-5917,1			
écart-type	20,6	821,8			

**Tableau A3:** Score des modèles de monomères calculées en utilisant 3MQ1\_A comme patron (chaîne A du dimère de Der p 5) et présentant une moyenne de z-score inférieure à -0,5. La plus basse énergie est en gras.

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
blot5_monomer <sub>1</sub> [3MQ1_A]	-124.0	-9957.0	-0.04	-1.90	-0.97
blot5_monomer <sub>2</sub> [3MQ1_A]	-111.0	<b>-10256.0</b>	0.31	-2.19	-0.94
blot5_monomer <sub>3</sub> [3MQ1_A]	-148.0	-9681.1	-0.68	-1.63	-1.16
blot5_monomer <sub>4</sub> [3MQ1_A]	<b>-235.0</b>	-9258.7	-3.01	-1.23	<b>-2.12</b>
blot5_monomer <sub>5</sub> [3MQ1_A]	-141.0	-8963.7	-0.49	-0.94	-0.72
blot5_monomer <sub>6</sub> [3MQ1_A]	-148.0	-8624.3	-0.68	-0.61	-0.65
blot5_monomer <sub>7</sub> [3MQ1_A]	-170.0	-8605.7	-1.27	-0.59	-0.93
blot5_monomer <sub>8</sub> [3MQ1_A]	-162.0	-8161.5	-1.05	-0.16	-0.61
Moyenne	-154.9	-9188.5			
écart-type	37.5	731.0			

	score		z-score		<z-score>
		Gromos	Anolea	Gromos	
derf5_monomer <sub>1</sub> [3MQ1_A]	-181.0	<b>-11283.3</b>	-0.64	-2.38	-1.51
derf5_monomer <sub>2</sub> [3MQ1_A]	<b>-232.0</b>	-10298.1	-2.20	-1.51	<b>-1.85</b>
derf5_monomer <sub>3</sub> [3MQ1_A]	-220.0	-9263.2	-1.83	-0.59	-1.21
derf5_monomer <sub>4</sub> [3MQ1_A]	-203.0	-9191.7	-1.32	-0.52	-0.92
derf5_monomer <sub>5</sub> [3MQ1_A]	-200.0	-10003.6	-1.22	-1.24	-1.23
Moyenne	-207.2	-10008.0			
écart-type	19.6	856.1			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
suim5_monomer <sub>1</sub> [3MQ1_A]	-144.0	-8572.4	-0.89	-1.30	-1.09
suim5_monomer <sub>2</sub> [3MQ1_A]	-151.0	<b>-8609.9</b>	-1.10	-1.33	-1.22
suim5_monomer <sub>3</sub> [3MQ1_A]	-173.0	-6819.5	-1.76	0.28	-0.74
suim5_monomer <sub>4</sub> [3MQ1_A]	<b>-186.0</b>	-6363.7	-2.15	0.69	-0.73
suim5_monomer <sub>5</sub> [3MQ1_A]	-153.0	-8848.1	-1.16	-1.55	<b>-1.35</b>
suim5_monomer <sub>6</sub> [3MQ1_A]	-116.0	-9061.4	-0.04	-1.74	-0.89
suim5_monomer <sub>7</sub> [3MQ1_A]	-140.0	-7978.3	-0.76	-0.77	-0.77
Moyenne	-151.9	-8036.2			
écart-type	22.7	1049.4			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
lepd5_monomer <sub>1</sub> [3MQ1_A]	-144.0	-7952.1	-1.12	-1.43	-1.27
lepd5_monomer <sub>2</sub> [3MQ1_A]	<b>-232.0</b>	<b>-8569.8</b>	-2.52	-2.13	<b>-2.33</b>
lepd5_monomer <sub>3</sub> [3MQ1_A]	-157.0	-7744.6	-0.18	-1.20	-0.69
lepd5_monomer <sub>4</sub> [3MQ1_A]	-181.0	-7679.7	-0.93	-1.12	-1.03
lepd5_monomer <sub>5</sub> [3MQ1_A]	-192.0	-6978.4	-1.27	-0.33	-0.80
lepd5_monomer <sub>6</sub> [3MQ1_A]	-217.0	-5789.8	-2.05	1.02	-0.52
Moyenne	-194.3	-7452.4			
écart-type	26.7	961.7			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
derp21_monomer <sub>1</sub> [3MQ1_A]	-181.0	-7847.3	-1.76	-0.79	-1.27
derp21_monomer <sub>2</sub> [3MQ1_A]	-85.0	-8540.0	0.07	-1.43	-0.68
derp21_monomer <sub>3</sub> [3MQ1_A]	<b>-189.0</b>	-8389.4	-1.91	-1.29	-1.60
derp21_monomer <sub>4</sub> [3MQ1_A]	-106.0	-7966.0	-0.33	-0.90	-0.61
derp21_monomer <sub>5</sub> [3MQ1_A]	-179.0	<b>-8627.2</b>	-1.72	-1.51	<b>-1.62</b>
derp21_monomer <sub>6</sub> [3MQ1_A]	-144.0	-8385.2	-1.06	-1.28	-1.17
derp21_monomer <sub>7</sub> [3MQ1_A]	-119.0	-7875.4	-0.58	-0.81	-0.69
derp21_monomer <sub>8</sub> [3MQ1_A]	-155.0	-6972.3	-1.26	0.03	-0.62
Moyenne	-144.8	-8075.4			
écart-type	38.3	539.3			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
derf21_monomer <sub>1</sub> [3MQ1_A]	-112.0	-6568.2	-0.66	-1.10	-0.88
derf21_monomer <sub>2</sub> [3MQ1_A]	-107.0	<b>-7215.5</b>	-0.52	-1.73	-1.13
derf21_monomer <sub>3</sub> [3MQ1_A]	-138.0	-6391.3	-1.35	-0.93	-1.14
derf21_monomer <sub>4</sub> [3MQ1_A]	<b>-188.0</b>	-7193.0	-2.69	-1.71	<b>-2.20</b>
derf21_monomer <sub>5</sub> [3MQ1_A]	-102.0	-6661.1	-0.39	-1.19	-0.79
derf21_monomer <sub>6</sub> [3MQ1_A]	-154.0	-6863.0	-1.78	-1.39	-1.58
derf21_monomer <sub>7</sub> [3MQ1_A]	-130.0	-6224.7	-1.14	-0.77	-0.96
Moyenne	-133.0	-6731.0			
écart-type	30.5	380.5			

	score		z-score		<z-score>
	Anolea	Gromos	Anolea	Gromos	
blot21_monomer <sub>1</sub> [3MQ1_A]	-197,0	-7087,3	-0,8	-1,0	-0,9
blot21_monomer <sub>2</sub> [3MQ1_A]	-204,0	<b>-7770,1</b>	-1,1	-1,9	-1,5
blot21_monomer <sub>3</sub> [3MQ1_A]	-179,0	-7161,1	-0,2	-1,1	-0,7
blot21_monomer <sub>4</sub> [3MQ1_A]	-226,0	-7748,4	-1,9	-1,9	<b>-1,9</b>
blot21_monomer <sub>5</sub> [3MQ1_A]	-205,0	-6313,9	-1,1	-0,1	-0,6
blot21_monomer <sub>6</sub> [3MQ1_A]	<b>-241,0</b>	-5803,9	-2,4	0,6	-0,9
blot21_monomer <sub>7</sub> [3MQ1_A]	-196,0	-7210,2	-0,8	-1,2	-1,0
Moyenne	-206,9	-7013,6			
écart-type	20,6	722,4			

**Tableau A4:** Valeurs de la fonction de score de ClusPro des meilleurs modèles prédits de la structure dimérique des allergènes en utilisant 3MQ1\_A comme patron (le meilleur score est en gras) et valeurs du rmsd obtenues après superpositions des modèles avec la structure 3MQ1\_AB

	Fonction de score de ClusPro	Rmsd (Å)
blot5_dimer <sub>1</sub>	-833	17,2
blot5_dimer <sub>2</sub>	-998	8,0
blot5_dimer <sub>3</sub>	-953	12,1
blot5_dimer <sub>4</sub>	-882	14,1
blot5_dimer <sub>5</sub>	-944	12,4
blot5_dimer <sub>6</sub>	<b>-1152</b>	11,6
blot5_dimer <sub>7</sub>	-914	20,2
blot5_dimer <sub>8</sub>	-1032	11,7

	Fonction de score de ClusPro	Rmsd (Å)
blot21_dimer <sub>1</sub>	-1539,9	2,46
blot21_dimer <sub>2</sub>	-1228	11,95
blot21_dimer <sub>3</sub>	-1276,4	2,28
blot21_dimer <sub>4</sub>	-1102,5	8,39
blot21_dimer <sub>5</sub>	-1461,1	2,89
blot21_dimer <sub>6</sub>	-1473,4	12,62
blot21_dimer <sub>7</sub>	<b>-1574</b>	8,86

	Fonction de score de ClusPro	Rmsd (Å)
derf5_dimer <sub>1</sub>	-1290	12,1
derf5_dimer <sub>2</sub>	-1187	12,2
derf5_dimer <sub>3</sub>	-1112	10,8
derf5_dimer <sub>4</sub>	-1186	12,4
derf5_dimer <sub>5</sub>	<b>-1353</b>	3,0

	Fonction de score de ClusPro	Rmsd (Å)
suim5_dimer <sub>1</sub>	-920	11,9
suim5_dimer <sub>2</sub>	-982	9,3
suim5_dimer <sub>3</sub>	-1094	15,2
suim5_dimer <sub>4</sub>	<b>-1168</b>	18,0
suim5_dimer <sub>5</sub>	-1116	2,9
suim5_dimer <sub>6</sub>	-1093	2,8
suim5_dimer <sub>7</sub>	-1013	6,9

	Fonction de score de ClusPro	Rmsd (Å)
lepd5_dimer <sub>1</sub>	-1033	16,5
lepd5_dimer <sub>2</sub>	-987	13,7
lepd5_dimer <sub>3</sub>	<b>-1190</b>	14,8
lepd5_dimer <sub>4</sub>	-1136	15,3
lepd5_dimer <sub>5</sub>	-1081	13,7
lepd5_dimer <sub>6</sub>	-1167	1,4

	Fonction de score de ClusPro	Rmsd (Å)
derp21_dimer <sub>1</sub>	-1178	13,8
derp21_dimer <sub>2</sub>	-1137	13,3
derp21_dimer <sub>3</sub>	-1206	12,7
derp21_dimer <sub>4</sub>	-868	15,6
derp21_dimer <sub>5</sub>	-1068	3,7
derp21_dimer <sub>6</sub>	<b>-1266</b>	13,0
derp21_dimer <sub>7</sub>	-1080	10,4
derp21_dimer <sub>8</sub>	-1108	13,2

	Fonction de score de ClusPro	Rmsd (Å)
derf21_dimer <sub>1</sub>	-950	13,2
derf21_dimer <sub>2</sub>	<b>-1252</b>	10,6
derf21_dimer <sub>3</sub>	-1030	8,8
derf21_dimer <sub>4</sub>	-1054	15,3
derf21_dimer <sub>5</sub>	-736	13,3
derf21_dimer <sub>6</sub>	-1133	11,3
derf21_dimer <sub>7</sub>	-874	11,0

### **Annexe III: Liste de communications et de publications effectuées durant ces années de thèse**

#### Communications orales

- **Khemili S.**, Hamadouche T. et Gilis D. 13-18 décembre 2010. "Modeling of the 3D structure of families 5 and 21 house dust mite allergens", *EMBO Practical Lecture course on Bioinformatics and Comparative Genome Analysis*, Institut Pasteur Tunis, **Tunisie**.

#### Communications affichées (Posters)

- **Khemili S.**, Lahiani S., Kwasigroch J.M., Hamadouche T., Galleni M. and Gilis D. Octobre 2012. "Prediction of B cell epitopes of the house dust mite allergen Der p 5", Immunoinformatics and Computational Immunology Workshop (ICIW2012), *Orlando, Floride (USA)*.
- **Khemili S.**, Kwasigroch J.M., Hamadouche T. et Gilis D. Septembre 2012. "Étude *in silico* de la structure 3D des allergènes majeurs d'acariens de poussière de maisons des familles 5 et 21: perspectives pour la conception de vaccins hypoallergéniques", 5ème Congrès de Médecine et de Chirurgie, *Tlemcen, Algérie*.
- **Khemili S.**, Kwasigroch J.M., Hamadouche T. et Gilis D. Mai 2012. "In silico dimerization propensity study of house dust mite allergens from families 5 and 21", Adjuvant2012 & Allergen Vaccines 2012 INTERNATIONAL CONGRESS, *Varadero, Cuba*.
- **Khemili S.**, T. Hamadouche et D. Gilis. Novembre 2010. "Modeling of the 3D structure of families 5 and 21 house dust mite allergens", *Communication affichée (Poster)*. LIX Bioinformatics Colloquium 2010, Ecole Polytechnique de Paris, **France**.
- **Khemili S.**, T. Hamadouche et D. Gilis. Octobre 2009. "Bioinformatic study of families 5 and 21 of house dust mite allergens", Journée doctorale SFMBBM (Structure et Fonction des Macromolécules Biologiques, Bioinformatique et Modélisation). Université de Namur, **Belgique**.
- **Khemili S.**, Akmouci S., Hamadouche T. et Gilis D. "Modélisation comparative d'un allergène majeur d'acarien (Der p 5)", Quatrièmes Journées Nationales de Biologie 23-24 Mai 2009. Université M'Hamed Bougara de Boumerdes, **Algérie**.

#### Article dans un journal international à comité de lecture

- **Khemili S.**, Kwasigroch J.M., Hamadouche T. et Gilis D. (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. *Journal of Biomolecular Structure & Dynamics*. Volume 29(4): 663-673.

#### Prix

Le prix du meilleur Poster m'a été décerné lors du 5ème Congrès de Médecine et de Chirurgie, *en Septembre 2012 à Tlemcen, Algérie*.

## Open Access Article

The authors, the publisher, and the right holders grant the right to use, reproduce, and disseminate the work in digital form to all users.



Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102  
Volume 29, Issue Number 4, (2012)  
©Adenine Press (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<sup>1\*</sup>  
Jean Marc Kwasigroch<sup>2</sup>  
Tarik Hamadouche<sup>1</sup>  
Dimitri Gilis<sup>2</sup>

## Abstract

Allergy represents an increasing threat 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

<sup>1</sup>Département de Biologie, Université M'Hamed Bougara de Boumerdes, Algeria.

<sup>2</sup>Unité de Bioinformatique génomique et structurale, Université Libre de Bruxelles, Brussels, Belgium.

\*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 putrescentiae*, 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  $\alpha$ -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  $\alpha$ -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 Blosom62 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.

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

where  $S_i$  is the score of the  $i$ -th model computed with Anolea or Gromos,  $\langle S \rangle$  is the average of the scores of the 30 models, and  $\sigma$  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,  $\langle z\text{-score}_i \rangle$ , 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<sub>1→8</sub>[3MQ1\_A]), 5 for Der f 5 (derf5\_monomer<sub>1→5</sub>[3MQ1\_A]), 7 for Sui m 5 (suim5\_monomer<sub>1→7</sub>[3MQ1\_A]), 6 for Lep d 5 (lepd5\_monomer<sub>1→6</sub>[3MQ1\_A]), 7 for Der f 21 (derf21\_monomer<sub>1→7</sub>[3MQ1\_A]) and 8 for Der p 21

(derp21\_monomer<sub>1→8</sub>[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<sub>1→10</sub>[2JMH]), 10 for Sui m 5 (suim5\_monomer<sub>1→10</sub>[2JMH]), 9 for Lep d 5 (lepd5\_monomer<sub>1→9</sub>[2JMH]), 8 for Der f 21 (derf21\_monomer<sub>1→8</sub>[2JMH]) and 9 for Der p 21 (derp21\_monomer<sub>1→9</sub>[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<sub>1→8</sub>, derf5\_dimer<sub>1→5</sub>, suim5\_dimer<sub>1→7</sub>, lepd5\_dimer<sub>1→6</sub>, derf21\_dimer<sub>1→7</sub>, derp21\_dimer<sub>1→8</sub>. 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<sup>α</sup> atoms and to compute the total rmsd deviation between the C<sup>α</sup>'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*,  $P_{X/Y}^i$ , has been computed according to the following equation:

$$P_{X/Y}^i = \frac{N_{X/Y}^i}{N_Y^i} \cdot 100\% \quad [2]$$

where  $N_{X/Y}^i$  is the number of common interactions of type *i* at the interfaces of the dimers *X* and *Y* and  $N_Y^i$  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

### ***Dimeric Structure of House Dust Mite Allergens***

PDB numbering (3MQ1)	34										40					45									
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	L	E	T	K	I	D	Q	V	E	K	G	A
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	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	D	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	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 than 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.

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

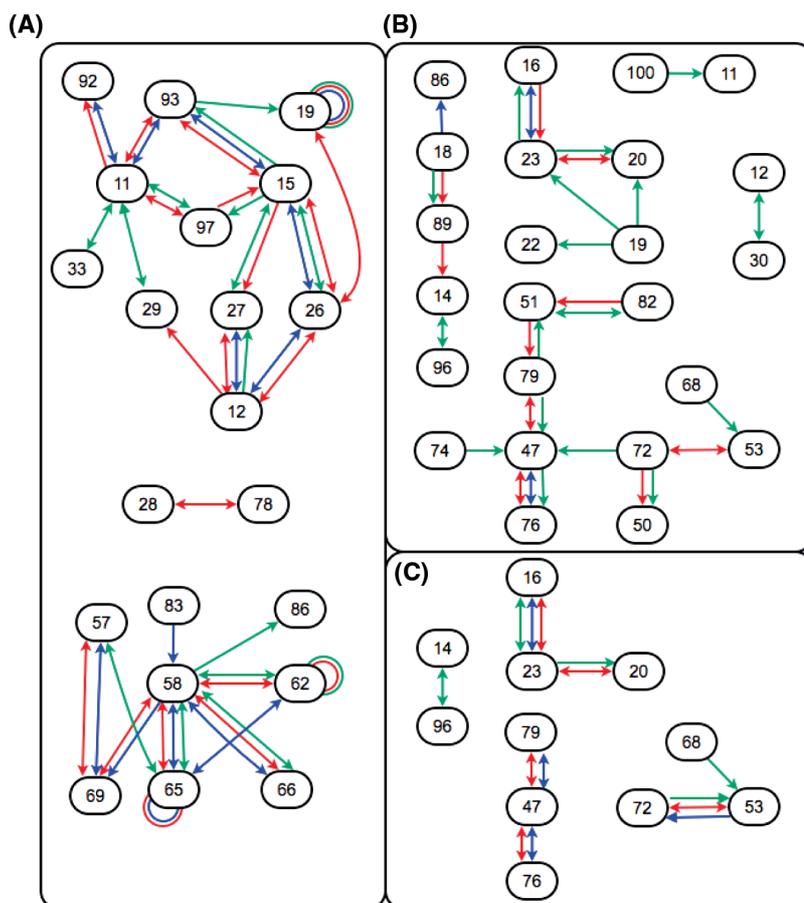
We used the ClusPro2.0 server (19) to predict the structure of the dimers. To test the ability of ClusPro to predict the dimeric structure of these allergens, we first used it to compute the structure of the Der p 5 dimer and compared this predicted structure to the X-ray structure of the dimer (chains 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<sup>α</sup> 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

## ***Dimeric Structure of House Dust Mite Allergens***



**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<sub>s</sub>) 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<sup>3</sup> in 3MQ1\_AB and of 2914 Å<sup>3</sup> in derp5\_dimer[3MQ1\_A]. Moreover, 83% of the

**Table IV**  
Percentage of common interface interactions between 3MQ1\_AB, derp5\_dimer[3MQ1\_A] and derf5\_dimer<sub>5</sub>.

	3MQ1_AB			derp5_dimer[3MQ1_A]			derf5_dimer <sub>5</sub>		
	$P_{3MQ1/X}^{\text{hydrophob}}$	$P_{3MQ1/X}^{\text{Hbond}}$	$P_{3MQ1/X}^{\text{salt-bridge}}$	$P_{derp5/X}^{\text{hydrophob}}$	$P_{derp5/X}^{\text{Hbond}}$	$P_{derp5/X}^{\text{salt-bridge}}$	$P_{derf5/X}^{\text{hydrophob}}$	$P_{derf5/X}^{\text{Hbond}}$	$P_{derf5/X}^{\text{salt-bridge}}$
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 <sub>5</sub>	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<sub>5</sub>). 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.  $P^{\text{hydrophob}}$ ,  $P^{\text{Hbond}}$  and  $P^{\text{salt-bridge}}$  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.

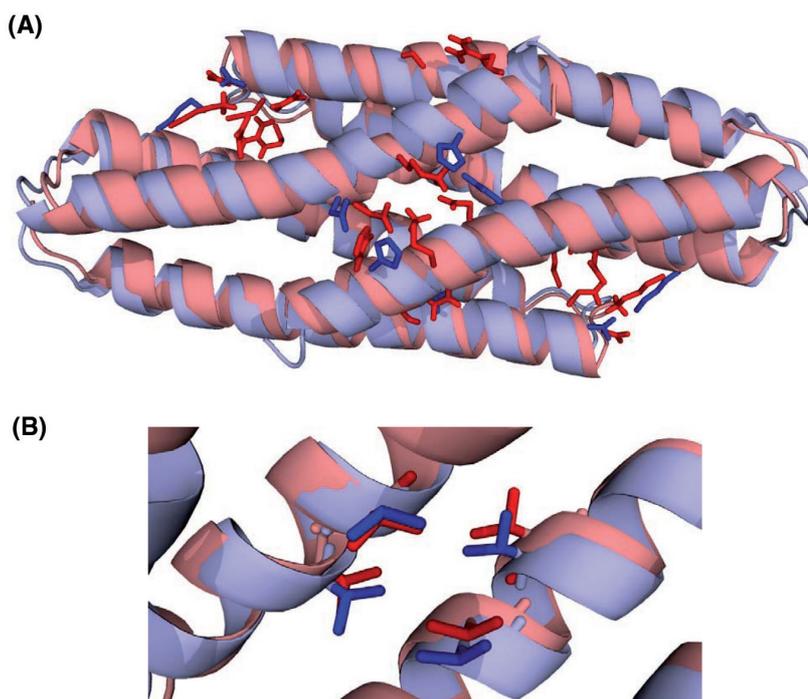
Der f 5 is the only allergen that presents a predicted dimeric structure similar to that of Der p 5 (see Table V, derf5\_dimer<sub>5</sub>). The rmsd after superimposition between Der p 5 (3MQ1\_AB) and derf5\_dimer<sub>5</sub> is equal to 3.0 Å; 48% of the interactions identified at the interface of derf5\_dimer<sub>5</sub> are common to those found in

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 <sub>6</sub>	-1152	>10
derf5_dimer <sub>5</sub>	-1353	3.0
suim5_dimer <sub>4</sub>	-1168	>10
lepd5_dimer <sub>3</sub>	-1190	>10
derp21_dimer <sub>6</sub>	-1266	>10
derf21_dimer <sub>2</sub>	-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.



**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.

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<sub>5</sub> to those found in derp5\_dimer[3MQ1\_A]. The volume of the cavity obtained by the assembly of both monomers in derf5\_dimer<sub>5</sub> is equal to 1979 Å<sup>3</sup>.

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\_dimer<sub>5</sub>, 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<sub>6</sub>[3MQ1\_A] and derf5\_monomer<sub>5</sub>[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<sub>1</sub>[3MQ1\_A]), which complicates the formation of an entangled dimer. The helix is more kinked in blot5\_monomer<sub>2</sub>[3MQ1\_A] (see Table S1b 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<sup>3</sup>.

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<sup>3</sup>.

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 a 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. Stęcklinger, 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. Bereman, 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

