

# **CHAPITRE 1: Introduction**

# 1. L'allergie

## Introduction

L'allergie est une hypersensibilité de type 1, c'est-à-dire médiée par les immunoglobulines de type E (IgE) (Coombs and Gell, 1975). Cette hypersensibilité est induite par certains antigènes appelés « allergènes ». Dans l'allergie, la réponse immunitaire induite possède toutes les caractéristiques d'une réponse humorale normale. L'allergène induit une réponse identique à celle produite par les autres antigènes conduisant à la création de plasmocytes et de cellules mémoires. Cependant, en plus des IgM et des IgG, les plasmocytes produisent des IgE qui sont normalement produits uniquement lors d'une infection parasitaire (Goldsby *et al.*, 2001).

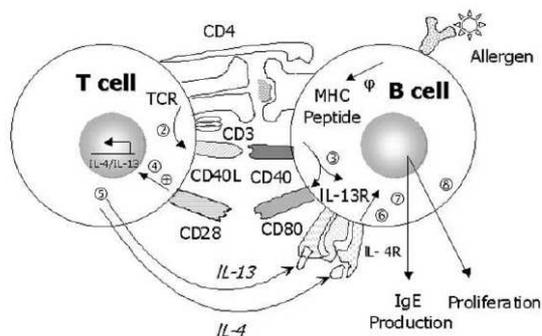
Certaines personnes présentent une prédisposition héréditaire, appelée « atopie », caractérisée par des troubles de la synthèse des IgE permettant à des antigènes non parasitaires de stimuler la production d'IgE (Goldsby *et al.*, 2001). Ces derniers possèdent une grande affinité pour les récepteurs de type FcεRI présents à la surface des mastocytes tissulaires et des basophiles du sang. Les mastocytes et les basophiles recouverts par les IgE sont alors dits « sensibilisés ». Lors d'une exposition ultérieure au même allergène, la fixation de l'allergène aux IgE présents à la surface des cellules sensibilisées provoque la dégranulation de ces dernières libérant des médiateurs pharmacologiques (leucotriènes, histamine) qui agissent sur les cellules avoisinantes et conduisent aux symptômes allergiques (rhinite allergique, asthme, dermatite).

La réponse IgE anormale des individus atopiques est, au moins en partie, d'ordre génétique et souvent familiale (Ruffilli and Bonini, 1997; Goldsby *et al.*, 2001). Les individus atopiques possèdent un taux anormalement élevé d'IgE circulantes et un nombre d'éosinophiles circulants supérieur à la normale (Goldsby *et al.*, 2001). La propension génétique aux réponses atopiques a été localisée dans divers loci chromosomiques. Un locus, situé sur le chromosome 5q, est lié à une région qui code pour une grande variété de cytokines, y compris l'IL-3, l'IL-4, l'IL-5, l'IL-9 l'IL-13 et le GM-CSF (Cookson *et al.*, 2000). Un second locus, situé sur le chromosome 11q, est lié à une région qui code pour une sous-unité du récepteur de haute affinité aux IgE (Cookson *et al.*, 1989). La plupart des réponses allergiques à IgE se produisent à la surface des muqueuses, en réponse à des allergènes qui pénètrent dans l'organisme soit par inhalation, soit par ingestion (Goldsby *et al.*, 2001).

## Base moléculaire de l'allergie

Lors de la réponse allergique, il y a une polarisation de la réponse immunitaire vers la production de lymphocytes T helper de type II (CD4<sup>+</sup> Th2) et une augmentation de production de cytokine de type II (IL-4, IL-5, IL-13) provoquant une augmentation de la production d'IgE et un développement des mastocytes, des basophiles et des éosinophiles (Miescher and Vogel, 2002) (Figure 1). Après capture et internalisation, l'allergène est protéolysé par une cellule présentatrice d'antigènes (cellule dendritique). Ce traitement, réalisé au niveau du protéasome, génère des peptides qui sont ensuite présentés à la surface de la cellule dans un contexte d'histocompatibilité de type II (CMH II). Ce complexe peut alors être reconnu par le récepteur complémentaire (TCR) d'un lymphocyte T helper naïf. En plus de cette interaction, les lymphocytes T ont besoin d'un second signal, appelé « signal de co-stimulation » qui résulte de l'interaction transitoire entre les molécules des cellules présentatrices

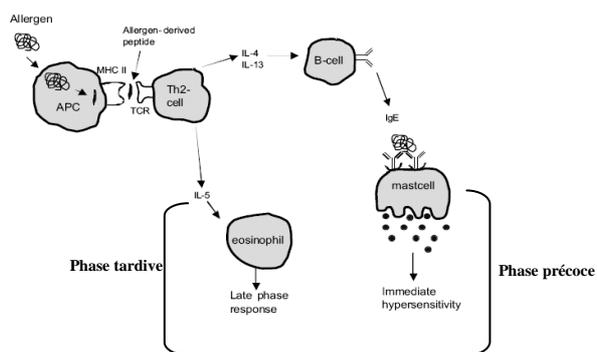
d'antigènes telles que CD80/CD86, DC-SIGN, CD40 avec des molécules de CD28, ICAM-3 et CD40L du lymphocyte T (Miescher and Vogel, 2002).



**FIGURE 1 : Régulation de la synthèse des IgE par les lymphocytes B et T (Miescher and Vogel, 2002)** 1 : la fixation de l'allergène à un IgM présent à la surface d'une cellule B (B-cell) conduit à son internalisation, sa dégradation au niveau du protéasome et sa présentation à la surface de la cellule dans un contexte d'histocompatibilité de type II (MHC peptide). 2 : la reconnaissance du complexe par le TCR d'une cellule T conduit à l'expression du CD40L. 3 : la fixation du CD40L au CD40 augmente l'expression du CD80. 4 : l'interaction CD80-CD28 stimule la production de cytokines (5) et la prolifération de lymphocytes T. 6 : la fixation de l'IL-13 et de l'IL-4 aux récepteurs présents à la surface des cellules B induit leur prolifération. 7 : « Switch isotypique » vers la production d'IgE. 8 : prolifération des lymphocytes B.

Cette co-stimulation permet le déclenchement de la différenciation et de l'expansion clonale des lymphocytes T helper qui se différencient alors en deux populations différentes selon le type de cytokines qu'ils produisent. Les Th1 qui produisent principalement de l'interféron- $\gamma$  (IFN- $\gamma$ ), de l'IL-2 et du TNF- $\beta$  et les Th2 qui produisent de l'interleukine de type 4, 5, 9, 10, 13 et de GM-CSF.

Les cytokines sécrétées par les Th2 stimulent la réponse allergique de différentes façons. L'IL-4 et l'IL-13 contribuent au phénomène allergique en stimulant le « switch isotypique » des lymphocytes B de la production d'IgM et d'IgG vers la production d'IgE (Del Prete *et al.*, 1988; Pene *et al.*, 1988). Les interleukines de type 3, 4 et 10 augmentent la production des mastocytes. L'IL-4 et l'IL-5 favorisent respectivement leur maturation, leur activation et l'accumulation des éosinophiles. En revanche, l'IL-2, l'INF- $\gamma$  et le TNF inhibent la production des IgE en agissant directement sur les lymphocytes B.



**FIGURE 2 : Vue simplifiée de la réponse allergique adaptée de (van Neerven *et al.*, 1996).** L'allergène est internalisé, processé et présenté à la surface d'une cellule présentatrice d'antigène (APC) dans un contexte d'histocompatibilité de type II (MHCII). L'interaction du complexe avec le récepteur complémentaire (TCR) d'un lymphocyte T helper (Th2 Cell) conduit celui-ci à produire des cytokines dont l'IL-4 et l'IL-13. Ces dernières induisent la production d'IgE par les lymphocytes B (B-cell) et sont responsables de la phase précoce. L'IL-5 induit la prolifération des éosinophiles et est responsable de la phase tardive.

Les immunoglobulines de type E, bien que présentes en faible quantité, jouent un rôle majeur dans la réponse allergique. Après production, celles-ci fixent le récepteur de haute affinité, FC $\epsilon$ RI, présent à la surface des mastocytes et des éosinophiles. La fixation de l'allergène provoque la libération de médiateurs préformés tels que l'histamine, l'IL-4, l'IL-5 et l'IL-13 conduisant à l'entrée en phase précoce du syndrome allergique (Janeway *et al.*, 1999) (Figure 2). Cette phase apparaît après quelques minutes et est caractérisée par des symptômes tels que des démangeaisons, des éternuements et des rhinites allergiques. La phase tardive de l'allergie apparaît dans les 6 à 24 heures suivantes et est

caractérisée par une augmentation des éosinophiles et des lymphocytes T, conduisant à des symptômes inflammatoires tels que l'hyperactivité bronchique (Mygind *et al.*, 1996).

### **L'allergie aux acariens**

L'allergie aux acariens est une pathologie chronique très répandue dans les pays industrialisés. A l'heure actuelle, on estime que plus de 20% de la population occidentale est sensibilisée aux acariens et que plus de 80% des asthmatiques développent des titres élevés en IgE spécifiques contre les allergènes de ces organismes (Thomas *et al.*, 2002). Ces réactions d'hypersensibilité immédiate de type 1 sont induites par différents types d'allergènes présents dans la poussière domestique et produits par différentes espèces d'acariens dont les deux prédominantes sont les espèces *Dermatophagoïdes pteronyssinus* et *Dermatophagoïdes farinae* (Voorhorst and Spieksma, 1967; Miyamoto *et al.*, 1968; Thomas *et al.*, 2002).

## **2. L'acarien *Dermatophagoïdes pteronyssinus* et ses allergènes**

### **L'acarien de poussière**

#### Morphologie

Les acariens sont des arthropodes appartenant à la classe des arachnides. Il existe environ 40 000 espèces différentes d'acariens (Nathanson, 1969). Les acariens possèdent une forme globulaire et une taille comprise entre 0,1 et 1 mm. Ils possèdent 4 paires de pattes et une cuticule striée. Leur durée de vie est comprise entre 2 et 3 mois et demi. Les acariens de poussière se divisent en deux groupes : les acariens de poussière de maison (House Dust Mite) qui se développent à l'intérieur des habitations et que l'on retrouve en association avec la poussière ménagère et les acariens de stockage qui se développent dans les milieux liés à l'agriculture.

Le cycle de vie des acariens est accompli en environ 30 jours et comprend plusieurs stades : oeuf, larve, nymphes et adulte. Leurs conditions de température et d'humidité optimales se situent respectivement entre 23 et 30°C et entre 80 et 90% d'humidité relative. La principale source de nourriture des acariens de poussière de maison est composée de squames de peau au contraire des acariens de stockage qui se nourrissent de plantes et de microorganismes. Les déjections des acariens de poussière représentent une puissante source d'allergènes. En effet, dans l'atmosphère, les fèces d'acariens contenant ces allergènes peuvent avoir la taille d'un grain de pollen et sont aisément mobilisables leur permettant ainsi de pénétrer facilement dans les voies respiratoires.

### **L'acarien *Dermatophagoïdes pteronyssinus***

#### Phylogénie et distribution



FIGURE 3 : L'acarien *Dermatophagoïdes pteronyssinus* (fr.wikipedia.org)

Bien que les deux espèces *Dermatophagoïdes pteronyssinus* (Figure 3) et *Dermatophagoïdes farinae* soient très proches et capables de croître dans des conditions similaires, leur répartition géographique

est différente. *D. pteronyssinus* est depuis longtemps considérée comme l'espèce prédominante en Europe et en Asie, au contraire des Etats-Unis où les deux espèces sont largement distribuées. Cependant, à l'heure actuelle, il semble que *D. farinae* devienne de plus en plus fréquente en Europe (Zock *et al.*, 2006).

### Les différents groupes d'allergènes d'acariens

Les acariens sont des organismes complexes qui peuvent produire des milliers de protéines différentes ainsi qu'une grande diversité de macromolécules. Les extraits d'acariens correspondent à des extractions liquides/solides d'une culture, épuisée ou non, contenant des acariens entiers, des nymphes, des pellettes fécales, des œufs, ainsi que du milieu de culture. Ces extraits sont capables d'induire à la fois une réactivité sur la peau de patients allergiques et de fixer leurs IgE dans des tests *in vitro* (Thomas *et al.*, 2002).

#### Nomenclature et mode de regroupement

La nomenclature systématique des allergènes est dictée par l'IUIS et contient trois parties séparées chacune par un espace. Les deux premières parties correspondent aux trois premières lettres du genre de l'organisme d'où l'allergène est isolé, suivi d'un espace, suivi de la première lettre de l'espèce. La troisième partie, désignée par un chiffre arabe, correspond à l'ordre chronologique d'identification de l'allergène. Dans les extraits d'acariens, on estime qu'environ 30 protéines différentes seraient capables d'induire la production d'IgE chez des patients allergiques aux acariens (Thomas *et al.*, 2002). Ces différentes protéines sont réparties en différents groupes en fonction de l'ordre chronologique de leur découverte et de leur fonction biologique (Thomas and Smith, 1998).

#### Groupes d'allergènes : description et prévalence

A l'heure actuelle, 15 groupes différents d'allergènes ont été identifiés et partiellement caractérisés chez *D. pteronyssinus*. La fonction biologique des groupes 1, 3, 4, 6, 9-11, 13-15 est connue alors que celle des groupes 2, 5, 7, 12, 21 et 23 reste inconnue. Un résumé de la fonction biologique de chaque groupe et de leur prévalence (% IgE) est présenté à la Table 1.

Groupe	Fonction	MM (kDa)	Espèce	IgE (%)
1	Protéase cystéine active	25	Dp, Df, Em	80-100
2	Inconnue (HE1)	14	Dp, Df, Em	80-100
3	Trypsine	25	Dp, Df, Em	16-100
4	$\alpha$ -amylase	57	Dp, Em	40-46
5	Inconnue	15	Dp	50-70
6	Chymotrypsine-like	25	Dp, Df,	40
7	Inconnue	25	Dp,Df,	50
8	Glutathione-S-	26	Dp	40
9	Collagénase	30	Dp	90
10	Tropomyosine	37	Dp, Df	50-95
11	Paramyosine	96	Df, Bt	80
12	Inconnue	14	Bt	50
13	Protéine fixant les acides	15	Bt	10-23
14	Vitellogenine/apolipop	177	Dp, Df, Em	90
15	Chitinase	62,5	Df	70
16	Gelsoline	55	Df	35
17	Protéine fixant le Ca <sup>2+</sup>	30	Df	35
18	Chitinase	60	Df	60
19	peptide Anti-microbien	7	Bt	10
20	Arginine kinase	-	Dp	15-44
21	Inconnue	14,7	Dp	56
23	Inconnue	14	Dp	-

**Table 1 : Les différents groupes d'allergènes d'acariens.** Adapté de (Thomas *et al.*, 2002) ([www.allergen.org](http://www.allergen.org))

## Der p 1 : l'allergène du groupe 1 de *Dermatophagoides pteronyssinus*

### Découverte

L'allergène Der p 1 fut le premier allergène découvert chez l'acarien *D. pteronyssinus* (Chapman and Platts-Mills, 1980) et le premier identifié par clonage de l'ADN complémentaire (Chua *et al.*, 1988). L'allergène Der p 1 (EC 3.4.22.65) est une glycoprotéine de 25 kDa, majoritaire dans les extraits d'acariens et capable d'induire la production d'IgE chez 80% des personnes sensibilisées aux acariens (Thomas and Smith, 1998).

### Classification, homologie et polymorphisme

Le gène de Der p 1 code pour une préproprotéine de 320 acides aminés comprenant une séquence signal de 18 acides aminés nécessaire pour la sécrétion, un propeptide de 80 acides aminés et une forme mature de 222 acides aminés correspondant à une protéase à cystéine active (Chua *et al.*, 1988) (Figure 4).

```

1  MKIVLATASLLALSAVYARPSSIKTFEEYKKAFNKSYATFEDEEAARKNFLESVKYVQSN
60  GGAINHLSDSLDEFKNRFLMSAEAFEHLKTQFDLNAETNACSSINGNAPAEIDLQMRVTV
120  TPIRMQGGCGSCWAFSGVAATESAYLAYRNQSLDLAEQELVDCASQHGCHGDTIPRGIEY
180  IQHNGVVQESYYRYVAREQSCRRPNAQRFGISNYCQIYPPNVNKIREALAQTHTSAIAVII
240  GIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNAQGVDYWIVRNSWDTNWGDNGYG
300  YFAANIDLMMIEEYPYVVIL

```

FIGURE 4 : Séquence de la protéine préproDer p 1 (PO8176). La séquence signal de 18 acides aminés et le propeptide de 80 acides aminés sont surlignés respectivement en noir et en gris. Les deux sites potentiel de N-glycosylation présents dans le propeptide (N16p) et dans la forme mature (N52) sont colorés en bleu. Les résidus de la triade catalytique C34-H170-N190 sont colorés en rouge.

Sur base de son homologie de séquence et de la conservation des résidus catalytiques, Der p 1 fut classé dans la famille des « papain-like » cystéine protéase (CA1) (Barrett *et al.*, 1998). En effet, proDer p 1 présente 14/ 26 % d'identité avec la propapaïne, 13/31 % avec la procaricaine et 12/27 % avec la cathepsine L (propeptide/forme mature) (Chua *et al.*, 1988). Le gène de Der p 1 est aussi présent chez d'autres espèces d'acariens très proches de *D. pteronyssinus*. Ainsi les allergènes Der f 1, Eur m 1 et Pso 1 ont été identifiés respectivement chez *D. farinae*, *Euroglyphus maynei* et *Psoroptes ovis* (Dilworth *et al.*, 1991). Ces protéines présentent respectivement de 85/82 %, 85/ 84 %, 79/ 64 % d'identité (propeptide/forme mature) ainsi qu'une prévalence vis-à-vis des IgE comprise entre 80 et 100%. Le gène de Der p 1 présente un polymorphisme au niveau d'au moins 12 résidus, ce qui le distingue de Der f 1 qui, lui, ne possède un polymorphisme qu'au niveau de 2 résidus (Smith *et al.*, 2001; Piboonpocanun *et al.*, 2006). Chez Der p 1, les deux isoformes les plus courantes contiennent la substitution de la valine 124 en alanine ou du glutamate 215 en lysine (Dilworth *et al.*, 1991; Smith *et al.*, 2001; Yasuhara *et al.*, 2001).

### Prévalence

La protéine Der p 1 a été identifiée comme étant un des allergènes majeurs de *Dermatophagoides pteronyssinus*, car cet allergène, en plus de sa grande concentration dans les extraits d'acariens, est capable de fixer des IgE chez tous les patients allergiques aux acariens. De plus, 50 à 70% des IgE produits en réponse aux extraits d'acariens sont dirigés contre Der p 1 (Chapman and Platts-Mills, 1980). Par la suite, des tests d'absorption ont montré que, selon les *sera*, la spécificité anti-D.

*pteronyssinus* est de 6 à 21% des IgE totaux (Lind, 1985). De plus, Van der Zee *et al.* montrèrent que, chez 60% des patients, 50 à 100 % des IgE produits en réponse aux extraits d'acariens sont dirigés contre Der p 1 et Der p 2 (van der Zee *et al.*, 1988). Des résultats similaires furent obtenus par Meyer (Meyer *et al.* 1994). A l'heure actuelle, on estime que Der p 1 est capable d'induire la production d'IgE chez 80 % des patients sensibilisés aux acariens (Tovey *et al.*, 1981; Chapman *et al.*, 1983; Krilis *et al.*, 1984; Meyer *et al.*, 1994).

### Structure

En 2005, la structure tridimensionnelle d'une forme doublement mutée (C34A/N52Q) du zymogène proDer p 1 a été résolue par diffraction des rayons X (Meno *et al.*, 2005) (PDB : 1XKG) (Figures 4A et 4B). L'analyse de cette structure confirme que Der p 1 appartient à la famille des «papain-like» protéases CA1. Cette structure montre que Der p 1 possède une structure globulaire à deux domaines délimitant un sillon catalytique à leur interface et maintenus ensemble par des interactions entre les extrémités N et C-terminales du zymogène (Figure 4C). Le premier domaine, appelé domaine R (Right), est formé par les résidus 197 à 299 dont la conformation est caractérisée par la présence d'hélices  $\alpha$  et de feuillets  $\beta$ . Ce domaine est traversé par les résidus 82 à 100 de l'extrémité N-terminale. Le second domaine, appelé domaine L (Left), est formé par les résidus 101 à 196 qui adoptent une conformation prédominante en hélices  $\alpha$ . Dans Der p 1, deux des trois ponts disulfure de la papaine sont conservés (C111-C151 et C145-C183). Cependant, un troisième pont disulfure, unique dans la famille CA1 et impliquant les résidus C84 et C197 est présent dans Der p 1 (Meno *et al.*, 2005).

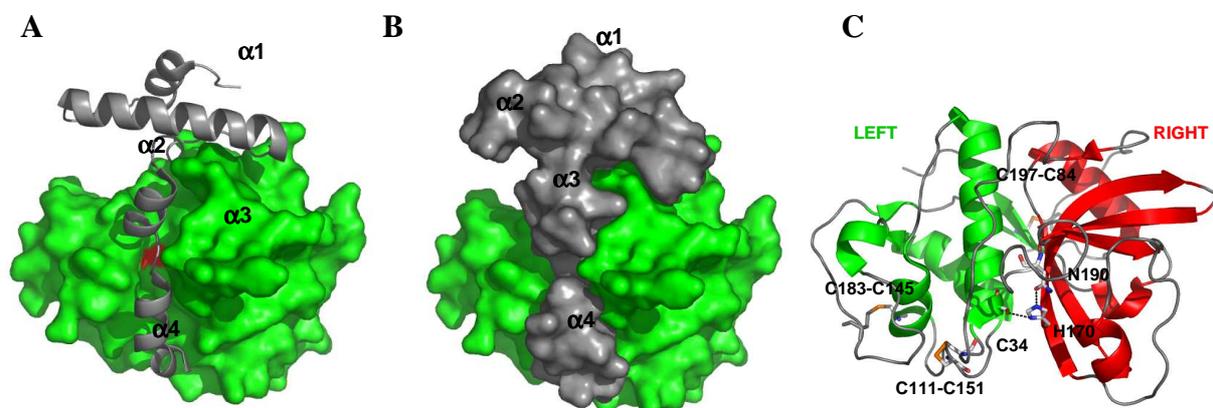


FIGURE 4 : **Structure du zymogène proDerp1, localisation du propeptide et des trois ponts disulfures** (Meno *et al.*, 2005). A et B: Structure du zymogène proDer p 1 (1XKG) et localisation des différentes hélices  $\alpha$  du propeptide. La protéase mature est colorée en vert et représentée en surface tandis que le propeptide est coloré en gris et représenté en cartoon (A) ou surface (B). Le site catalytique est coloré en rouge. C: structure de la protéase mature et localisation des trois ponts disulfure et de la triade catalytique (C34-H170-N190).

En 2006, la structure de Der p 1 a été résolue, par diffraction des rayons X, à partir d'une forme mature recombinante provenant du zymogène proDer p 1 (N52Q) préalablement activé par un traitement à pH 4 (de Halleux *et al.*, 2006) (PDB : 2AS8). L'analyse de cette structure a confirmé les résultats obtenus par Meno *et al.* L'alignement de ces deux structures montre qu'il n'y pas de différences structurales notables entre la structure de la forme mature libre et celle obtenue à partir du zymogène (Figure 5A). Ces résultats indiquent que la structure du domaine catalytique n'est pas modifiée après le processus d'activation ou par la présence du propeptide. Cependant, la structure de Der p 1 obtenue par De Halleux *et al.* montre que Der p 1 est capable de dimériser à pH neutre et

basique (Figures 5B et 5C) et de fixer un ion magnésium. Cette dimérisation, observée dans la maille cristalline et par tamis moléculaire, mais absente dans la structure du zymogène, serait rendue possible par des interactions masquées en présence du propeptide et impliquant trois segments (69-75, 146-148 et 163-168), portant chacun des résidus polaires et aromatiques.

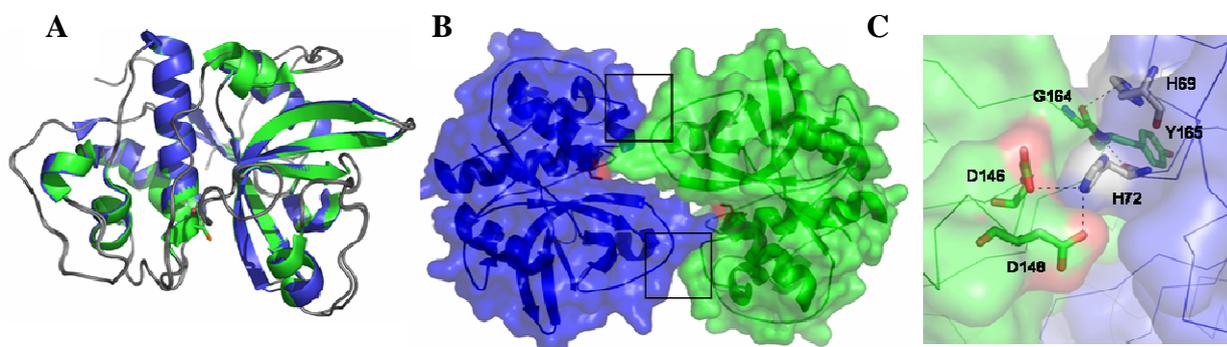


FIGURE 5 : **Dimérisation de la protéase Der p 1 mature** (de Halleux *et al.*, 2006). A: superposition de la structure du domaine catalytique du zymogène (IXKG) en bleu et de la protéase mature (2AS8) en vert. B: implication du sillon catalytique dans la dimérisation de la protéase mature et localisation (encadrés) des résidus impliqués dans la stabilisation de l'interface de dimérisation. C: représentation et localisation des résidus impliqués dans la stabilisation de l'interface entre les deux monomères.

La dimérisation de Der p 1, bien qu'unique dans le clan CA1, a déjà pu être mise en évidence lors de la cristallisation de la Ginger protéase II (famille CA1), cette dernière adoptant une structure homotétramérique (Choi *et al.*, 1999). Deux résidus histidine (His69 et His72) seraient des résidus clés dans cette dimérisation de Der p 1, expliquant ainsi que cette interaction est dépendante du pH. Le résidu His 72 serait un résidu crucial dans la stabilisation de l'interface. En effet, l'histidine 72 d'un monomère forme un pont hydrogène avec l'Asp146 et l'Asp148 de l'autre monomère et son carbonyle forme un pont hydrogène avec l'azote de la chaîne principale de la tyrosine 165. Cette dernière s'insère dans une poche hydrophobe délimitée par les résidus Ile76, Ile141, Tyr169, Tyr216 et Tyr218 de son partenaire. L'histidine 69, quant à elle, semble moins importante, mais son azote Nε2 forme un pont hydrogène avec le carbonyle de la Gly164. Enfin, l'asparagine 168, située au milieu de la surface d'interaction interagit avec son équivalent sur l'autre monomère et la Gln166 forme un pont hydrogène avec la tyrosine 169 de l'autre partenaire. De plus, l'analyse de la structure de Der p 1 a montré la présence d'un site de fixation d'un ion magnésium, également observée par Meno *et al.*, mais cet ion ne serait pas essentiel à l'activité catalytique car Der p 1 doit être activée par un agent réducteur (DTT) et un agent chélateur des ions bivalents (EDTA) afin de régénérer le thiol libre de la cystéine active (Meno *et al.*, 2006).

#### Activité et fonction

L'analyse de la structure tridimensionnelle montre que le site actif de Der p 1, situé à l'interface entre les domaines L (vert) et R (rouge) (Figure 4C), est formé d'une paire ionique thiolate-imidazolium conservée chez les « papain-like » et comprenant les chaînes latérales des résidus Cys34, His170 et Asn190. Dans cette triade, la cystéine active est portée par le domaine Left tandis que l'histidine et l'asparagine sont portées par le domaine Right. Bien qu'à l'heure actuelle aucune fonction physiologique précise n'ait pu lui être attribuée, Thomas *et al.* ont pu mettre en évidence par immunodétection une concentration élevée en Der p 1 au niveau de l'épithélium de la partie postérieure de

l'estomac de l'acarien, ce qui suggère une fonction digestive (Thomas *et al.*, 1991). Cette hypothèse est renforcée par la présence d'une concentration élevée en Der p 1 dans les pellettes fécales (50% des protéines totales) de l'acarien. Cependant, la présence d'un peptide signal semble indiquer que le zymogène Der p 1 pourrait être soit sécrété dans la lumière du système digestif, soit adressé à un compartiment cellulaire avant d'être activé.

### Spécificité

Les protéases à cystéine active de la famille CA1 sont majoritairement des endopeptidases. Chez celles-ci, la fixation du substrat est assurée par sept sous-sites S4-S1 et S1'-S3', répartis sur les deux côtés du sillon catalytique, chacun d'entre eux étant capable de stabiliser une chaîne latérale du substrat (Schechter and Berger, 1967). Dans cette famille, le déterminant le plus important de la spécificité est le sous-site S2 qui affiche, dans la majorité des cas, une grande préférence pour les larges chaînes latérales hydrophobes comme celle de la phénylalanine en position P2. Au contraire, le sous-site S1 est moins sélectif dans cette famille ce qui est vraisemblablement dû à une absence de poche de spécificité définie. Dans la famille CA1, peu de choses sont connues concernant la spécificité des sous-sites S' mais il semble que leur spécificité soit relativement étendue.

Chez Der p 1, la spécificité protéolytique a pu être bien caractérisée par des expériences de protéolyse et par des études structurales. Dans ces différentes expériences, Der p 1 est décrite comme étant une protéase possédant une très grande spécificité au niveau du résidu en position P2 du substrat. En effet, Schulz a étudié la protéolyse de 10 substrats tripeptidiques fluorescents de séquences différentes par la protéase Der p 1 naturelle purifiée à partir de cultures épuisées d'acariens (Schulz *et al.*, 1998b). Cette étude a montré qu'un seul de ces substrats, le Boc-Gln-Ala-Arg-AMC, pouvait être hydrolysé par Der p 1 ( $K_m = 280 \mu\text{M}$ ,  $k_{cat}/K_m = 4,6 \cdot 10^3 \text{ M}^{-1} \text{ s}^{-1}$ ). Elle a également démontré l'importance de la présence d'une alanine en position P2 et d'un résidu chargé positivement en position P1, ce qui semble la distinguer des autres « papain-like » protéases.

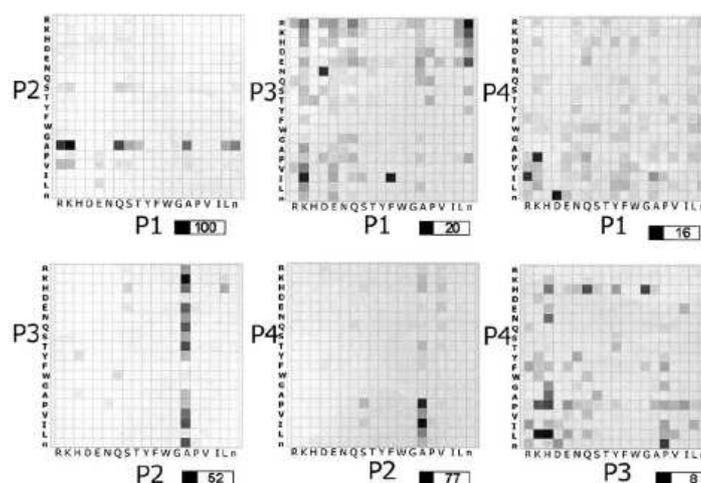


FIGURE 6: Spécificité protéolytique de la protéase Der p 1 naturelle déterminée à l'aide d'une librairie de tétrapeptides fluorescents (Harris *et al.*, 2004). Etude de l'impact de la nature et de la position des différents résidus du substrat et de la relation entre les différentes positions suivant la nomenclature de Schechter and Berger (Schechter and Berger, 1967). Le numéro situé en dessous de chaque sous-librairie représente le pourcentage d'activité considérant le signal le plus important comme 100% (activité maximale).

D'autre part, Harris a pu confirmer ces résultats en étudiant la protéolyse d'une banque combinatoire de 4000 tétrapeptides fluorescents immobilisés sur des damiers (Figure 6). Ces résultats confirment la préférence de Der p 1 pour un résidu alanine en position P2. Cette étude montre également la préférence de Der p 1 pour un acide aminé basique en position P1, polaire en P3, ainsi que pour la présence d'un acide aminé aliphatique tel qu'une leucine, une isoleucine, une proline ou une valine en position P4 (Harris *et al.*, 2004).

Les bases moléculaires de cette spécificité ont pu être expliquées en partie par une étude approfondie des résidus longeant le site catalytique de Der p 1 et par leur comparaison avec ceux de la papaïne. En effet, cette comparaison a pu mettre en évidence plusieurs différences, majoritairement dues à la présence de nombreuses substitutions au niveau des résidus déterminant les différents sous-sites et par la position différente de l'hélice  $\alpha$  et de la boucle portant les résidus 224 à 248. Ainsi, ces différences conduiraient chez Der p 1 à l'absence de sous-site S1 et à la possibilité de stabiliser un acide aminé à large chaîne latérale hydrophobe dans le sous-site S1'. De plus, bien que la chaîne latérale en position P1 du substrat semble devoir pointer en dehors du site catalytique, une possible interaction avec l'asparagine 248, située à l'extérieur du sillon catalytique pourrait expliquer la préférence de Der p 1 pour un acide aminé polaire et/ou basique en position 1 (Meno *et al.*, 2005).

#### Modifications post-traductionnelles

##### Glycosylation

L'analyse de la séquence de l'ADN complémentaire indique la présence de deux sites potentiels de N-glycosylation (NXS/T) dans le précurseur de Der p 1 (proDerp1). Le premier site serait situé dans le propeptide au niveau du résidu Asn16p (-64 numérotation forme mature) dans la boucle entre les hélices alpha 1 et 2 et le second dans la séquence de la forme mature portée par le résidu Asn52 (Chua *et al.*, 1988).

##### Protéolyse

Comme toutes les protéases à cystéine active de la famille des «papain-like», Der p 1 est synthétisée sous forme inactive afin d'éviter la protéolyse non contrôlée des cellules productrices. Dans la famille CA1, cette inactivation est rendue possible par la présence d'une extension N-terminale, appelée propeptide ou prosegment, qui couvre le sillon catalytique et diminue l'accessibilité du site catalytique. Cependant, ce mode de régulation nécessite une activation ultérieure correspondant à la perte du propeptide et à la libération de l'activité enzymatique. Cette activation peut être intra ou intermoléculaire en une ou plusieurs étapes selon les zymogènes.

#### Implication de l'activité enzymatique dans la réponse allergique

L'activité enzymatique de Der p1 joue un rôle important dans le développement de la réponse allergique (Kikuchi *et al.*, 2006). Plusieurs études suggèrent que Der p 1 pourrait causer des dommages directs à l'épithélium respiratoire. En effet, il a été montré que cette cystéine protéase peut cliver des protéines d'adhésion des jonctions serrées telles que l'occludine et la claudine (Wan *et al.*, 1999). De ce fait, Der p 1 augmenterait la perméabilité épithéliale et donc, faciliterait le mouvement transépithélial des allergènes. En plus du clivage de ces protéines, l'activité protéolytique de Der p 1

est capable d'inactiver un inhibiteur des protéases à sérine appelé l' $\alpha$ 1 antitrypsine (Kalsheker *et al.*, 1996). Cet inhibiteur, naturellement sécrété par les cellules épithéliales, est impliqué dans la régulation homéostatique de l'épithélium respiratoire. Par ce clivage, Der p 1 pourrait réduire la défense du tissu épithélial contre des protéases et augmenterait donc la vulnérabilité du tissu (Sharma *et al.*, 2003). De plus, cet allergène est capable de cliver l'A1-Pi ( $\alpha$ -1-proteinase inhibitor), la SLPI (secretory leukocyte proteinase inhibitor) et l'élafine qui contribuent significativement à l'activité anti-élastase dans les poumons (Brown *et al.*, 2003).

L'allergène Der p 1 est capable de cliver spécifiquement le récepteur CD23, récepteur de faible affinité aux IgE, présent à la surface des lymphocytes B (Schulz *et al.*, 1995). Ce clivage se produit à deux sites distincts dans le récepteur : un premier entre les résidus Glu298 et Ser299 et un second entre les résidus Ser155 et Ser156 (Schulz *et al.*, 1997). Le récepteur CD23 joue un rôle important dans la réponse immunitaire humorale. En effet, son activation par les IgE conduit à un signal « feedback » qui supprime la synthèse locale des IgE par les cellules B. Par ailleurs, il apparaît que, *in vitro*, la forme soluble découlant du clivage du récepteur CD23 augmente la synthèse d'IgE. Des études montrent que les lymphocytes T, conditionnés au préalable avec cet allergène, induiraient une augmentation de la synthèse des IgE par les lymphocytes B. L'activité protéolytique de Der p 1 dans l'allergie semble ainsi être liée à une dérégulation de la synthèse des IgE (Gough *et al.*, 1999).

De plus, Der p 1 est aussi capable de cliver le récepteur CD25 des cellules T. Ce clivage se produit dans la chaîne  $\alpha$  du récepteur de haute affinité à l'interleukine-2 et conduit à une diminution de la prolifération des lymphocytes T et de la production de IFN- $\gamma$  ; l'IL-2 et l'IFN- $\gamma$  étant considérés comme deux facteurs de croissance autocrine des Th1. Ainsi, le clivage du CD25 par Der p 1 tendrait à diminuer la population de Th1 et donc à faire augmenter celle de Th2 (Schulz *et al.*, 1998a).

Der p 1 peut aussi cliver les protéines surfactantes A et D des poumons (Deb *et al.*, 2007). Ces deux protéines, capables de fixer les carbohydrates en présence d'ions  $Ca^{2+}$ , joueraient un rôle important dans la lutte anti-bactérienne ainsi que dans la réponse allergique en interagissant avec une grande variété d'aéro-allergènes. En effet, ces protéines pourraient se fixer aux allergènes eux-mêmes ou à la place de ceux-ci sur les IgE présents à la surface des mastocytes, empêchant ainsi la dégranulation de ces derniers. La dégradation des surfactants A et D par Der p 1 conduirait ainsi à une diminution des défenses antibactériennes mais pourrait aussi augmenter la réponse et les symptômes allergiques.

L'activité protéolytique de Der p 1 peut favoriser la voie Th2 via un clivage des molécules DC-SIGN et DC-SIGNR (Furmonaviciene *et al.*, 2007). Ces deux protéines appartenant aux lectines de type C sont impliquées dans la reconnaissance des pathogènes. Le clivage de DC-SIGN par Der p 1 conduirait à une diminution de ICAM-3 et du ligand de DC-SIGN qui est normalement exprimé à la surface des lymphocytes T naïfs et qui est impliqué dans la signalisation de la voie Th1.

Enfin, l'activité enzymatique de Der p 1 peut aussi conditionner les cellules dendritiques à stimuler la différenciation des lymphocytes T naïfs en lymphocytes Th2. Der p 1 clive le récepteur CD40 présent à la surface des cellules dendritiques, conduisant à une diminution de la production d'IL-12. Ces cellules dendritiques, conditionnées par la protéine Der p 1, induiraient ainsi chez des lymphocytes T

naïfs une augmentation de la production d'IL-4 et une diminution de la synthèse d'IFN- $\gamma$  (Comoy *et al.*, 1998; Ghaemmaghami and Shakib, 2002).

Finalement, la comparaison des réponses induites par des protéines Der p 1 recombinantes inhibées par du E-64 ou dénaturées par la chaleur avec celles induites par la protéase active montre que les protéines dépourvues d'activité protéolytique induisent une production d'IgE totaux, d'IgE et d'IgG spécifiques de Der p 1 plus faibles que celle obtenue avec l'allergène Der p 1 actif (Kikuchi *et al.*, 2006).

Un résumé des principales implications de l'activité enzymatique de Der p 1 dans le déclenchement de la réponse allergique est présenté à la figure 7.

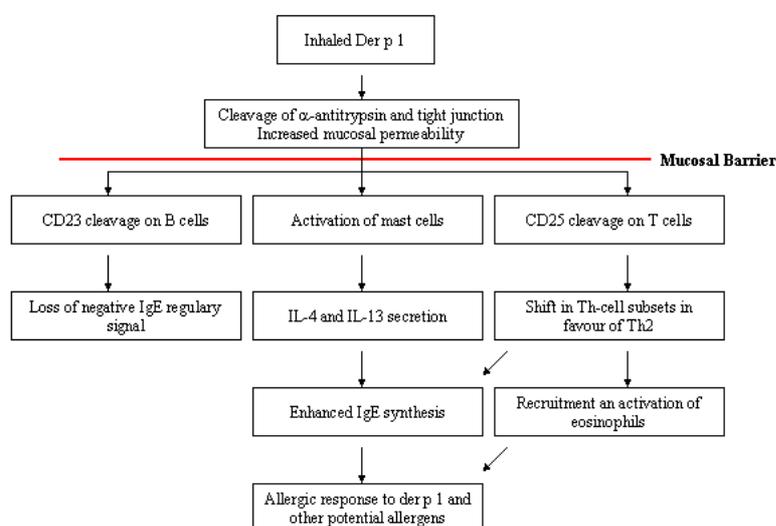


FIGURE 7: Implication de l'activité protéolytique de l'allergène Der p 1 dans la perméabilité pulmonaire et dans le déclenchement de la réponse allergique (Shakib *et al.*, 1998).

#### Source naturelle

L'allergène Der p 1 fut pour la première fois purifié à partir d'extraits d'acariens par Chapman en 1980 (Chapman and Platts-Mills, 1980). Dans ces extraits, Der p 1 est la protéine majoritaire dans les extraits d'acariens. Dans sa forme naturelle, l'allergène Der p 1 est glycosylé et présente une extrémité N-terminale unique caractérisée par la séquence TNA-. La nature exacte de sa glycosylation reste inconnue et sujette à discussion. Celle-ci pourrait faire intervenir une N-glycosylation de l'asparagine 52 comportant des mannoses comme suggéré par la prise en charge de l'allergène par le récepteur à mannose lors de sa phagocytose par les cellules dendritiques (Deslee *et al.*, 2002) ou des O-glycosylations comportant une structure du type  $\beta(1-3)$  N-acétylgalactosamine comme suggéré par une analyse réalisée à l'aide de différentes lectines (van Oort *et al.*, 2002).

#### Intérêts des allergènes recombinants

Les extraits naturels utilisés en diagnostic ou en désensibilisation sont généralement obtenus à partir de sources allergéniques complexes et dont la composition peut varier en fonction de leur origine, des procédés d'extraction, de purification et des conditions de stockage. En diagnostic, quelle que soit la méthode utilisée, l'utilisation des extraits naturels ne permet pas de préciser à quel allergène le patient

est sensibilisé ni de connaître la proportion d'IgE spécifiquement dirigées contre les différents allergènes contenus dans les extraits (Metz-Fabre *et al.*, 2006). En thérapie, l'utilisation d'extraits comporte le risque de sensibiliser le patient contre un allergène pour lequel il n'est pas sensible. Face à ces problèmes, les scientifiques ont, depuis une vingtaine d'années, tenté de produire des allergènes de manière recombinante afin d'obtenir ces derniers en grande quantité et avec un haut niveau de pureté. Ainsi, l'allergène Der p 1 fut le premier à avoir été cloné en 1988 (Chua *et al.*, 1988). En 2005, on comptait plus de 500 allergènes clonés provenant de différentes sources telles que le bouleau (Bet v 1, 2, 4), les graminées (Phl p 1-11), le latex (Hev b 1-11) et le chat (Fel d 1). L'obtention de ces formes recombinantes a ainsi permis d'améliorer la connaissance moléculaire des allergènes et de simplifier le diagnostic. L'expression recombinante de l'allergène Der p 1 a, par exemple, permis de montrer que cet allergène ne possédait qu'une activité de protéase à cystéine active et pas une double activité cystéine/sérine active comme il avait été proposé au terme de plusieurs études et de suggérer que cette activité de protéase à sérine active provienne d'une contamination par une trace d'un autre allergène présent dans l'extrait de départ (Takai *et al.*, 2005).

Enfin, le développement de systèmes d'expression d'allergènes recombinants permet d'envisager la production de formes modifiées d'un allergène appelées « hypoallergène » dans lesquelles les épitopes reconnus par les IgE sont éliminés tandis que les épitopes T et ceux reconnus par les IgG sont conservés, permettant l'induction IgG neutralisant (Valenta, 2002). Ces hypoallergènes peuvent être ensuite utilisés à des fins thérapeutiques, que ce soit pour traiter des sensibilités individuelles et spécifiques ou pour réduire les risques anaphylactiques liés à l'utilisation d'une immunothérapie basée sur l'utilisation d'allergènes spécifiques.

#### Essai d'expression recombinante de Der p 1

Comme la protéase Der p 1 est naturellement produite dans l'intestin de l'acarien, puis libérée dans l'environnement par les matières fécales, les quantités de Der p 1 naturelle disponibles sont extrêmement faibles. De plus, Der p 1 est contaminée par la présence d'autres allergènes tels que les protéases à sérine active Der p 3, 6 ou 9 (Takai *et al.*, 2005). Dès lors, l'obtention de grandes quantités de formes recombinantes de l'allergène présentant un grand degré de pureté et possédant toutes les caractéristiques de l'allergène naturel, que ce soit en terme d'activité enzymatique ou de reconnaissance par les IgE de patients allergiques, s'avère être une étape indispensable pour une meilleure compréhension de son rôle dans le développement de l'allergie, pour le diagnostic ainsi que dans le développement de thérapies vaccinales basées sur l'administration de vaccins hypoallergéniques. Durant la dernière décennie, l'expression recombinante de Der p 1 dans différents organismes hôtes a été expérimentée par différentes équipes avec plus ou moins de succès.

Le premier essai d'expression recombinante de la forme mature de l'allergène Der p 1 a été réalisé par Thomas *et al.* en bactérie *E. coli* sous forme d'une protéine de fusion (Thomas *et al.*, 1988). Cependant, dans cette étude, l'allergène était faiblement exprimé et majoritairement sous forme insoluble conduisant à une réactivité IgE jusqu'à 50% plus faible que celle mesurée pour l'allergène naturel. Par la suite, Chua *et al.* ont réalisé l'expression de Der p 1 en levure *S. cerevisiae* (Chua *et al.*, 1992). Dans ce cas, bien que la protéine exprimée présentait une grande reconnaissance par les IgE, des différences de réactivité aux IgE persistaient, comparées à l'allergène naturel.

La première tentative d'expression recombinante de Der p 1 sous forme de zymogène proDer p 1, a été décrite par Takahashi (Takahashi *et al.*, 2000). Dans cette étude, le zymogène proDer p 1 était produit sous forme insoluble (corps d'inclusion) dans le cytoplasme de *E. coli*. Cependant, malgré une resolubilisation complète, le zymogène exprimé n'a pu être activé en protéase Der p 1 et a montré encore une fois une reconnaissance vis-à-vis des IgE de patients allergiques bien moins importante que celle de l'allergène naturel. La même année, Best a démontré, pour la première fois, qu'il était possible de produire un zymogène de l'allergène proDer f 1, l'homologue de proDer p 1, correctement replié et reconnu par des IgE de patients allergiques en utilisant la levure méthylotrophe *Pichia pastoris* comme système d'expression (Best *et al.*, 2000). De manière inattendue, dans cette étude, le zymogène sécrété était mûri en une protéase active. D'autres essais de production du précurseur proDer p 1 ont été réalisés par Jacquet *et al.* en cellules de drosophile (Jacquet *et al.*, 2000) et en cellules de mammifères CHO (Massaer *et al.*, 2001). Dans ces deux systèmes, le zymogène était sécrété dans le milieu de culture et présentait une reconnaissance aux IgE équivalente à celle que l'allergène naturel indiquant un repliement correct de la protéase. Cependant, cette équipe n'a jamais pu mûrir ces zymogènes jusqu'à l'apparition d'une activité enzymatique. Enfin, des tentatives d'expression du précurseur proDer p 1 ont été tentées dans la levure méthylotrophe *Pichia pastoris* par plusieurs équipes (Jacquet *et al.*, 2002; Takai *et al.*, 2002; van Oort *et al.*, 2002). Ces essais ont démontré que le zymogène obtenu était activable par un traitement acide et que la protéase obtenue au terme de cette maturation était hyper-glycosylée et présentait une reconnaissance aux IgE équivalente à celle de l'allergène naturel.

### **Les autres allergènes d'acariens à activité peptidase**

Parmi les protéines présentes dans les extraits d'acariens de *D. pteronyssinus*, au moins trois autres allergènes possèdent une activité protéolytique. Ces protéases nommées Der p 3, Der p 6 et Der p 9 appartiennent respectivement aux groupes 3, 6 et 9. Ces protéines, bien que présentes en faible quantité dans les extraits d'acariens, ont pu être purifiées et partiellement caractérisées. Leur ADN complémentaire a pu être isolé et séquencé et leurs caractéristiques immunologiques sont partiellement connues.

#### **Der p 3 : allergène à activité « trypsin-like »**

L'allergène Der p 3 est une protéase à sérine active présente dans les d'extraits d'acariens et purifiée pour la première fois par Stewart en 1992 (Stewart *et al.*, 1992). Sur base de son homologie de séquence et de son activité protéolytique, Der p 3 (EC 3.4.21.-) a été classé dans la famille « trypsin-like » sérine protéase (S1). En effet, Der p 3 possède 45 % d'identité avec la trypsine bovine et 47% avec celle du saumon. Elle montre une grande préférence pour les substrats peptidiques possédant un résidu arginine ou lysine en position P1. L'analyse de son ADN complémentaire a montré que, comme pour Der p 1, Der p 3 est synthétisé sous forme d'une préprotéine de 261 acides aminés formée d'une séquence signal de 18 acides aminés, d'un propeptide de 11 acides aminés et d'une protéase mature de 232 acides aminés. En outre, le propeptide de Der p 3 présente une séquence très particulière « NPILPASP NAT » contenant un motif PxxPxxP et une extrémité C-terminale (PNAT-) unique et très éloignée de la séquence DDDDK caractéristique des « trypsin-like » protéases. De plus,

cette extrémité C-terminale contient un site potentiel de N-glycosylation (NXS/T) qui pourrait conduire à l'ajout de chaînes glycaniques sur l'asparagine en position -3.

Bien que son caractère d'allergène majeur soit controversé, Smith *et al.* ont reporté une prévalence contre le groupe 3 comprise entre 70 et 80% (Smith *et al.*, 1994). Tout comme pour Der p 1, la présence de Der p 3 dans les fèces d'acariens semble indiquer que cette protéase pourrait être, elle aussi, impliquée dans la digestion de l'acarien. Cette hypothèse semble renforcée par la fonction et la localisation de cette famille d'enzymes chez d'autres organismes. En effet, chez les mammifères, la trypsine est une enzyme digestive localisée dans l'intestin. Chez ces organismes, le mécanisme d'activation de la trypsine conduisant à la perte du propeptide et à la libération de l'activité enzymatique est connu. Le zymogène de la trypsine, appelé trypsinogène, est synthétisé par les cellules pancréatiques, puis activé dans l'intestin par l'entérokinase qui reconnaît spécifiquement et clive après le motif polyaspartyl-lysine du trypsinogène. Une fois activée, la trypsine peut ensuite reconnaître le motif DDDDK d'un autre trypsinogène et le cliver. Cependant, pour Der p 3, les particularités du propeptide et l'absence de motif DDDDK suggèrent que le mécanisme d'activation de proDer p 3 pourrait être unique dans la famille des « trypsin-like » protéases.

Enfin, l'activité protéolytique de Der p 3 joue aussi un rôle dans le déclenchement de l'allergie. En effet, Sun *et al.* ont montré que Der p 3 était capable de cliver le récepteur PAR-2 (protease activated receptor) de l'épithélium respiratoire pulmonaire (Sun *et al.*, 2001). Ce clivage, réalisé en aval du résidu Arg36, induirait une cascade signalétique qui conduirait au relargage de cytokines proinflammatoires telles que le GM-CSF (Granulocyte Macrophage Colony-Stimulating Factor), l'éotaxine ou l'IL-8 par les cellules épithéliales. De plus, Maruo a pu montrer que Der f 3, l'homologue de Der p 3 produit par *D. farinae* est capable de cliver les protéines C3 et C5 du complément et de générer ainsi les anaphylatoxines proinflammatoires C3a et C5a qui pourront induire notamment la dégranulation des mastocytes (Maruo *et al.*, 1997).

#### Der p 6 : allergène à activité « chymotrypsin-like »

L'allergène Der p 6 a été pour la première fois identifié en 1993 par Yasueda *et al.* dans des extraits d'acariens (Yasueda *et al.*, 1993). Cet allergène possède une activité protéolytique caractéristique des « chymotrypsin-like » et n'est pas sérologiquement lié au groupe 3. En 1996, Bennett a isolé l'ADN complémentaire de Der p 6 et confirmé son appartenance à la famille des « chymotrypsin-like » protéases S1 (EC 3.4.21.-) (Bennett and Thomas, 1996). Der p 6 est une protéase à sérine active (Ser184-His45-Asp91) de 231 acides aminés possédant les trois ponts disulfures et tous les motifs caractéristiques de cette famille. Der p 6 possède 75% d'identité avec son homologue Der f 6 et seulement 37% d'identité avec Der p 3, ce qui pourrait expliquer l'absence de réaction croisée entre les deux groupes. Tout comme pour les protéases Der p 1 et Der p 3, l'allergène Der p 6 est synthétisé sous forme d'un zymogène, appelé proDer p 6, précédé d'un signal de sécrétion de 16 acides aminés. Ce zymogène proDer p 6 contient un propeptide de 32 acides aminés, plus long que les propeptides habituellement retrouvés dans cette famille et son extrémité C-terminale ne se termine pas par un résidu arginine mais par une séquence (PIAG-) tout à fait unique dans cette famille. Enfin, le propeptide de Der p 6 possède lui aussi un site potentiel de N-glycosylation (NXS/T) sur l'asparagine en position -7.

Der p 9 : allergène à activité « collagenase-like »

L'allergène Der p 9 est une protéase à sérine active qui se distingue des protéases à sérine active Der p 3 et Der p 6 par son activité collagénase. Der p 9 fut isolé à partir d'extraits d'acariens et caractérisé pour la première fois en 1996 par King (King *et al.*, 1996). L'analyse de sa prévalence par ces mêmes auteurs a montré que Der p 9 est reconnu par 92% des *sera* de patients allergiques aux acariens.

### 3. Les zymogènes de protéases à cystéine active de la famille 1 « Papain-like »

#### Introduction

Les protéases sont divisées en sept groupes en fonction de leur mécanisme catalytique (Barrett *et al.*, 1998) (<http://merops.sanger.ac.uk>). Ces groupes correspondent aux protéases à acide aspartique, les métalloprotéases, les protéases à sérine, les protéases à cystéine, les protéases à thréonine, les protéases à acide glutamique et celles dont le mécanisme catalytique est inconnu. Les protéases à cystéine active se répartissent en quatre-vingt familles dont les six majeures sont : les papain-like, les calpain-like, les clostripaine-like, les protéases à cystéine active streptococcales, les protéases à cystéine active virales et les caspases.

Les protéases à cystéine active de la famille 1 «papain-like» (CA1) sont largement distribuées dans le vivant que ce soit chez les animaux, les végétaux, les virus ou les procaryotes (Rawlings and Barrett, 2000). Cette famille regroupe à la fois des endo et des exopeptidases qui sont impliquées dans une large gamme de processus protéolytiques spécifiques ou non. Cette famille de protéases est la plus grande des 80 familles de protéases à cystéine active. En 2008, pas moins de 1904 membres de cette famille étaient répertoriés dans la banque de données de protéases MEROPS.

#### Structure du domaine catalytique des protéases à cystéine active de la famille 1 (CA1)

La structure et la fonction de nombreuses protéases de la famille CA1 sont connues. La structure tridimensionnelle de la papaïne, le prototype de cette famille, a été résolue en 1968 par Drenth *et al.* (Drenth *et al.*, 1968). Depuis, les structures d'autres enzymes tels que la caricaine (Katerelos *et al.*, 1996), les cathepsines L (Fujishima *et al.*, 1997), K (McGrath *et al.*, 1997; Sivaraman *et al.*, 1999) et B (Musil *et al.*, 1991) ont été résolues. Ces structures ont révélé un arrangement tridimensionnel conservé chez tous les membres reflétant l'identité de séquence élevée. Dans ces structures, on peut distinguer deux domaines (Figure 8A).

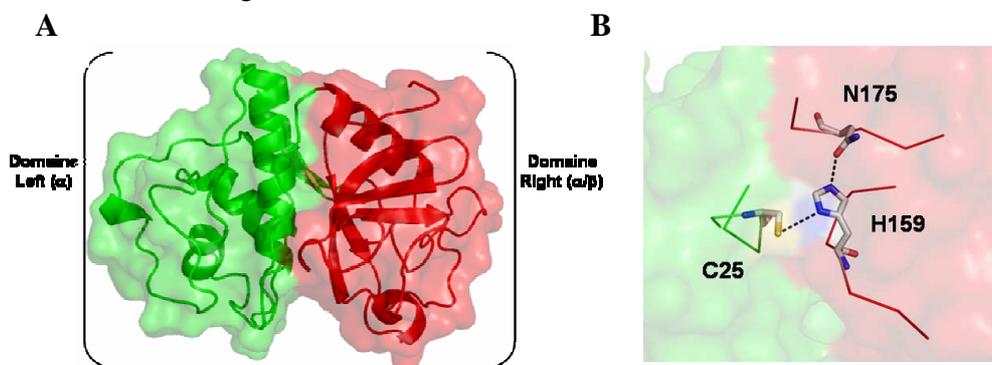


FIGURE 8: Structure de la papaïne (PDB: 9PAP), localisation des deux domaines et de la triade catalytique. A: Le domaine Left, formé majoritairement d'hélices  $\alpha$  est représenté en vert tandis que le domaine Right, formé d'hélices  $\alpha$  et de brins  $\beta$  est représenté en rouge. B: localisation de la triade C25-H159-N175 à l'interface entre les deux domaines.

Un domaine du N-terminal, appelé domaine Left, formé majoritairement d'hélices  $\alpha$  et dont la signature caractéristique est une grande hélice centrale longue de 30 acides aminés et un domaine C-terminal, appelé domaine Right, formé à la fois d'hélices  $\alpha$  et de brins  $\beta$ . Ces deux domaines interagissent entre eux sur une large surface et délimitent le sillon catalytique en forme de « V » contenant la triade catalytique Cys-His-Asn (Figure 8B).

### **Fonction et localisation cellulaire**

Les fonctions assurées par les protéases à cystéine active de la famille CA1 peuvent être très différentes et peuvent varier selon le type d'organisme. Chez les plantes, les protéases de type CA1 sont utilisées, par exemple, pour mobiliser les protéines de stockage dans les graines après l'induction de la germination (Schlereth *et al.*, 2001). La plupart des protéases à cystéine des parasites sont extracellulaires et aident le parasite pour son adhérence, pour l'acquisition de nutriments, pour sortir des cystes ou pour échapper au système immunitaire de l'hôte (Sajid and McKerrow, 2002). Chez les mammifères, les protéases à cystéine sont considérées comme lysosomiales mais peuvent parfois être sécrétées. Certaines se retrouvent dans tous les tissus comme les cathepsines B, C, H, L, O et pourraient être impliquées dans le maintien cellulaire alors que d'autres comme les cathepsines S, K, V, F, X et W ne se retrouvent que dans certains organes suggérant une fonction spécifique (Wiederanders, 2003).

### **Le zymogène : forme inactive et régulée de la protéase.**

L'activité protéolytique des protéases représente un danger potentiel pour les cellules qui les produisent. La régulation de l'activité enzymatique des protéases est donc essentielle car une protéolyse à un mauvais moment et/ou à un mauvais endroit pourrait être létale pour les cellules et les tissus producteurs (Neurath, 1984). Pour remédier à ce problème, plusieurs stratégies de régulation de l'activité des protéases ont été développées dans le vivant (Twining, 1994; Chapman *et al.*, 1997). Ces stratégies sont basées sur des principes très différents tels que la régulation de l'expression des gènes, les modifications post-traductionnelles, la production d'inhibiteurs spécifiques, l'autolyse, la compartimentation ou encore l'inhibition par le substrat ou le produit. Dans la majorité des cas, ces modes de régulation ne sont pas exclusifs mais utilisés en combinaison afin de limiter au maximum toute protéolyse non désirée.

Une autre stratégie largement utilisée dans différentes familles de protéases, telles que les sérine protéases (trypsinogène, prosubtilisine), les protéases à acides aspartiques (pepsinogène), les métalloprotéases (procarboxipeptidase A2) et les protéases à cystéine active (procathepsine L) consiste à produire la protéase sous une forme inactive, appelée zymogène, dans laquelle l'activité du domaine catalytique est inhibée par la présence d'une extension N-terminale de longueur variable appelée propeptide (Khan and James, 1998). Dans ces zymogènes, le propeptide longe le sillon catalytique et agit en tant qu'inhibiteur en limitant l'accessibilité du site catalytique et/ou en induisant des changements de position des résidus impliqués dans la catalyse ou dans la reconnaissance du substrat. L'activité protéolytique de l'enzyme pouvant ensuite être libérée, dans des conditions adéquates, lors d'un processus dit « d'activation » durant lequel le propeptide est enlevé par protéolyse intra ou intermoléculaire selon le zymogène considéré. Une des fonctions des propeptides étant d'inhiber l'activité protéolytique du domaine catalytique, il est intéressant de noter que la synthèse du

propeptide précède celle du domaine catalytique afin de limiter les protéolyses non désirées au cours de la synthèse et donc que le propeptide soit localisé à l'extrémité N-terminale de la protéase.

### **Propeptides de zymogènes de la famille CA1**

Toutes les protéases à cystéine active de la famille CA1 sont synthétisées sous forme d'un zymogène précédé d'un peptide signal. Les propeptides des protéases de la famille CA1 possèdent une taille comprise entre 38 et 251 résidus selon le zymogène et dans une moindre mesure selon l'espèce considérée (Dickinson, 2002). Dans les deux extrêmes, la cathepsine X possède un propeptide très court de 38 ou 41 acides aminés selon le site de clivage considéré, alors que celui de la dipeptidyl peptidase I est très grand et compte 206 résidus. Les propeptides sont très importants pour le domaine catalytique et joueraient un rôle primordial dans l'inhibition, le repliement et l'adressage des zymogènes.

### **Fonction du propeptide dans la famille CA1**

#### Propeptide et inhibition

Dans les zymogènes, le propeptide est un élément de régulation clé de l'activité de la protéase car en inhibant l'activité protéolytique, il empêche toute dégradation des cellules et des tissus producteurs (Reddy *et al.*, 1995). La première inhibition d'une protéase à cystéine active par un propeptide fut observée par Fox en 1992 (Fox *et al.*, 1992), en étudiant l'interaction entre la cathepsine B de rat et un peptide synthétique de 55 résidus correspondant aux résidus 7 à 62 de son propeptide. Cette inhibition était dépendante du pH et reliée à une dégradation du propeptide par la protéase à pH acide. Par la suite, Taylor a réalisé pour la première fois, en 1995, l'expression indépendante des propeptides de la papaïne et la papaya protease IV (PPIV) en *E. coli* et a pu mettre en évidence la spécificité de ces deux propeptides sur base de leur potentiel inhibiteur envers la papaïne, la caricaine, la chymopapaïne et la PPIV (Taylor *et al.*, 1995). D'autres études (Volkel *et al.*, 1996; Maubach *et al.*, 1997; Visal *et al.*, 1998; Billington *et al.*, 2000; Guay *et al.*, 2000) ont confirmé, par la suite, que l'inhibition des protéases à cystéine active de la famille CA1 par leur propeptide est une loi générale et que le propeptide représente un moyen de régulation de l'activité de ces protéases. Ces nombreux travaux, en particulier ceux réalisés sur les propeptides de cathepsines, ont permis de montrer que les propeptides de la famille CA1 inhibent leurs protéases respectivement avec une constante d'affinité de l'ordre du nanomolaire ou subnanomolaire comme démontré pour les cathepsines L, K et S (Guay *et al.* 2000) (Maubach *et al.* 1997) (Carmona *et al.*, 1996).

Les bases structurales de cette inhibition furent élucidées avec l'obtention des premières structures de zymogènes de protéases à cystéine active CA1 tels que la procathepsine B (Cygler *et al.*, 1996; Turk *et al.*, 1997) et la procathepsine L (Coulombe *et al.*, 1996). Ces structures révélèrent le mode d'inhibition utilisé par le propeptide et la place occupée par celui-ci dans le zymogène. Dans toutes les structures, le propeptide couvre le sillon catalytique et bloque l'accessibilité du site catalytique dans une orientation non productive par rapport au substrat, ce qui l'empêche d'être dégradé par la protéase. De plus, dans cette orientation, la partie N-terminale du propeptide recouvre les sous-sites S' de fixation au substrat tandis que la partie C-terminale bloque les sous-sites S, ce qui empêche toute reconnaissance du substrat (Schechter and Berger, 1967). Ce mode de fixation et d'inhibition est conservé dans tous les zymogènes de protéases à cystéine active de la famille CA1 résolus à ce jour.

### Propeptide et repliement

Toute l'information nécessaire à une protéine pour adopter sa structure tridimensionnelle et donc sa fonction biologique est contenue dans sa séquence (Anfinsen, 1973). Toutefois, lors du repliement, certaines protéines appelées «chaperonne moléculaire» peuvent aider la protéine à acquérir sa structure tridimensionnelle en créant un environnement favorable au repliement, en évitant son aggrégation, en favorisant la formation de ponts disulfure ou l'isomérisation cis-trans des liaisons X-Pro. Cependant, plusieurs observations semblent être en conflit avec cette loi universelle. En effet, pour de nombreuses protéases de différentes familles, leurs expressions sous forme de protéase mature ne conduit pas à l'obtention de protéases actives ce qui suggère que le propeptide pourrait être nécessaire au repliement en apportant spécifiquement l'information manquante pour le repliement de la protéase auquel il est associé. Cette hypothèse largement décrite pour la subtilisine (Shinde *et al.*, 1993) et la protéase  $\alpha$ -lytique (Peters *et al.*, 1998) indique que le propeptide de ces protéases jouerait un rôle de « chaperonne intramoléculaire » en diminuant l'énergie d'activation nécessaire au passage d'une forme intermédiaire de type « molten globule » vers l'état natif. Chez les protéases à cystéine active, ce rôle de chaperonne intramoléculaire fut tout d'abord proposé pour expliquer le mauvais repliement, observé *in vitro*, de formes tronquées de la procathepsine L à partir de l'extrémité N-terminale (Smith and Gottesman, 1989). Cette hypothèse fut ensuite confirmée par la nécessité que le propeptide de la cathepsine L soit présent afin de pouvoir synthétiser *in vivo* de la cathepsine L correctement repliée (Tao *et al.*, 1994; Ogino *et al.*, 1999). Cependant, il semble que cette fonction du propeptide n'est pas une loi générale chez tous les propeptides de la famille CA1 ou que l'entiereté du propeptide n'est pas toujours nécessaire. En effet, dans certaines tumeurs malignes, la cathepsine B est produite sous forme active et correctement repliée à partir d'un RNAm tronqué dans lequel le deuxième et le troisième exon sont manquants, ce dernier exon faisant partie du propeptide (Mehtani *et al.*, 1998). Enfin, il semble, au vu d'expériences de « refolding » de la cathepsine S en présence de propeptide libre que la liaison covalente entre le propeptide et la protéase mature ne soit pas essentielle à la fonction de chaperonne du propeptide (Pietschmann *et al.*, 2002).

### Propeptide et adressage

Enfin, la présence de séquences de reconnaissance ou de sites de glycosylation dans les propeptides pourrait être essentielle pour la sécrétion et l'adressage de certains zymogènes à l'intérieur de la cellule. Dans la famille CA1, l'adressage des zymogènes de protéases lysosomales peut être réalisé par deux voies possibles : la voie dépendante du mannose-6-phosphate et la voie indépendante.

#### Voie dépendante du mannose-6-phosphate

Dans la voie dépendante du mannose-6-phosphate, les signaux nécessaires à l'adressage vers les lysosomes ou les vacuoles acides peuvent être présents dans le propeptide ou dans la protéase mature. Cette voie est catalysée par une N-acétyl-glucosamine-1-phosphate transférase qui utilise un UDP-N-acétyl-glucosamine pour phosphoryler le mannose d'une N-glycosylation, la protéine manno-phosphorylée pouvant ensuite être reconnue par un récepteur spécifique et transportée vers le lysosome. La sélectivité de cette voie semble être structurale et liée à la présence d'au moins deux résidus lysine éloignés de 34 Å sur la surface de la protéine (Baranski *et al.*, 1991; Baranski *et al.*, 1992; Wiederanders, 2003). Dans la cathepsine L, ce motif est situé dans le propeptide et la mutation

d'une seule de ces deux lysines est suffisante pour empêcher l'adressage lysosomal (Cuozzo *et al.*, 1995; Cuozzo *et al.*, 1998). Cependant, dans la famille CA1, ce motif n'est pas toujours localisé dans le propeptide. En effet, chez la cathepsine B, ce motif est situé dans la forme mature. La glycosylation, elle aussi, peut être localisée uniquement sur le propeptide comme pour le cathepsine S ou sur la protéase mature selon les enzymes (Wiederanders *et al.*, 1992).

#### Voie indépendante du mannose-6-phosphate

Cette voie repose sur la présence d'une séquence conservée de 9 résidus (WKAMHNRLY) localisée près de l'extrémité N-terminale du propeptide (McIntyre *et al.*, 1994; Huete-Perez *et al.*, 1999). Ce motif ancestral semble être conservé chez la plupart des protéases de la famille CA1, que ce soit chez les plantes ou chez les animaux mais est absent chez les cathepsines B, F et X. Ce motif serait reconnu par un récepteur intramembranaire de 43 kDa permettant l'adressage vers les vacuoles acides ou lysosomales (McIntyre and Erickson, 1993).

### Classification en sous-familles et propeptides

En plus de leurs caractéristiques structurales et enzymatiques qui ont conduit à leur regroupement au sein des « papain-like » protéases, les protéases à cystéine active de la famille CA1 peuvent être classées en sous-familles. Cette classification, introduite par Karrer en 1993 est basée sur la longueur, la structure et la présence de motifs caractéristiques dans les propeptides de ces protéases (Karrer *et al.*, 1993). Dans cette classification, les zymogènes de la famille CA1 ont été subdivisés en deux sous-familles : les protéases dont le propeptide comporte le motif « ERFNIN » ou sous-famille « cathepsin L ou L-like » et les protéases dont le propeptides ne comporte pas de motif « ERFNIN » ou sous-famille « cathepsin B ou B-like ».

#### Propeptide de la sous-famille L

##### Caractéristiques structurales

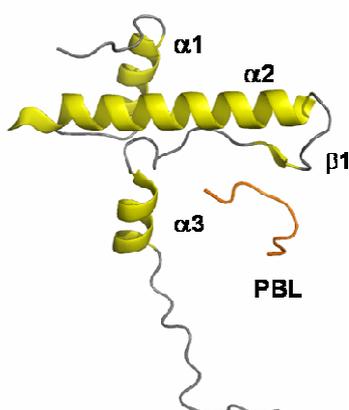


FIGURE 9: Structure du propeptide de la cathepsine L. Propeptide (jaune) et PBL (orange)

Les propeptides des protéases de la sous-famille L-like possèdent une taille d'environ 100 acides aminés et représentent donc environ un tiers de la taille du zymogène. Ces propeptides possèdent un repliement identique caractérisé par trois parties distinctes : un domaine globulaire N-terminal qui fixe la partie distale de l'enzyme, un élément de fixation au site catalytique qui bloque l'accessibilité du site catalytique dans une orientation non productive et un segment flexible C-terminal qui longe le sillon catalytique jusqu'à l'extrémité N-terminale de la protéase mature. Dans ces propeptides, le domaine globulaire N-terminal est formé par l'interaction entre les hélices  $\alpha 1$ ,  $\alpha 2$  et le brin  $\beta 1$ . Ce dernier est dirigé de manière antiparallèle par rapport à l'hélice  $\alpha 2$  et interagit avec le domaine catalytique au niveau d'une région appelée Propeptide Binding Loop (PBL) (Figure 9). Les deux hélices, pour leur part, forment entre elles un angle compris entre  $60^\circ$  et  $90^\circ$ . Ce domaine globulaire n'est pas directement impliqué dans l'inhibition mais déterminerait l'affinité et la spécificité entre le propeptide et la protéase mature (Carmona *et al.*, 1996) (Guay *et al.*, 2000). La partie C-terminale du

propeptide, quant à elle, s'imbrique tout le long du sillon catalytique et couvre les sous-sites de reconnaissance du substrat et bloque l'accessibilité au substrat.

#### Le motif ERFNIN

Le motif «ERFNIN» fut identifié pour la première fois par Karrer en 1993 sur base de l'alignement de séquences de 20 zymogènes de protéase à cystéine active de la famille CA1 (Karrer *et al.*, 1993). Cet alignement mit en évidence la présence d'un motif conservé, dans 15 des 20 séquences analysées, qui s'étend sur 20 acides aminés et consiste en une série de résidus conservés interrompus par une série de résidus variables selon la séquence  $EX_3RX_2(V/I)FX_2NX_3IX_3N$ . Son appellation, sous le nom de motif «ERFNIN», correspond à la séquence en code à une lettre de l'ensemble des résidus conservés. Sur base de la séparation des résidus conservés par des résidus variables, Karrer proposa que ce motif s'étende sur une face d'une hélice  $\alpha$  et puisse constituer une unité fonctionnelle. Plus tard, la résolution de la structure de plusieurs zymogènes de la sous-famille L, dont la caricaine, confirma la localisation du motif «ERFNIN» le long de la seconde hélice du propeptide (Figures 10A et 10B).

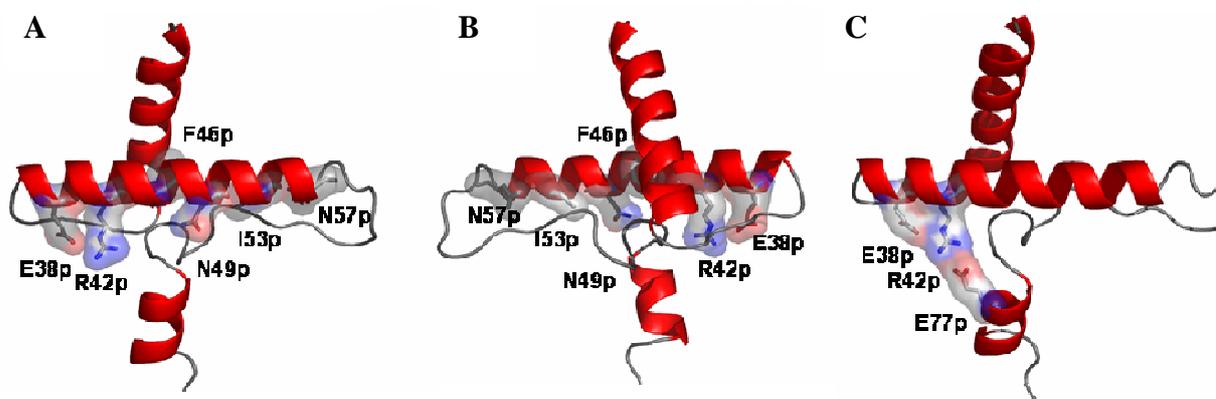


FIGURE 10: Localisation du motif ERFNIN dans le propeptide de la caricaine. A et B: mise en évidence et localisation des différents résidus impliqués dans le motif stabilisateur ERFNIN sur la face avant (A) et arrière (B) du propeptide de la caricaine. C: mise en évidence de la triade Glu38p-Arg42p-Glu77p reliant les hélices  $\alpha$  2 et 3.

De plus, la structure de la procaricaine montre qu'une partie des chaînes latérales des résidus du motif «ERFNIN» pointent vers l'intérieur du domaine globulaire et réalisent des interactions électrostatiques, aromatiques et des liaisons hydrogène avec les chaînes latérales des résidus de l'hélices  $\alpha$  1 et 3 ainsi qu'avec le brin  $\beta$  jouant potentiellement un rôle important dans l'architecture et la stabilité de ce domaine. Ainsi, la chaîne latérale du glutamate (E) du motif ERFNIN forme un pont salin avec l'arginine (R) de ce même motif, cette dernière formant de son côté un second pont salin avec un second glutamate situé sur l'hélice  $\alpha$ 3 (Figure 10C). Cette triade Glu-Arg-Glu permettrait ainsi d'orienter l'hélice  $\alpha$ 3 dans le sillon catalytique et stabiliserait l'association entre ces deux hélices. La phénylalanine (F), quant à elle, est impliquée dans la formation du cœur hydrophobe du domaine globulaire tandis que la première asparagine (N) est impliquée dans la formation d'un pont hydrogène avec la forme mature. Enfin, la chaîne latérale de l'isoleucine (I) pointe vers le cœur hydrophobe, tandis que la seconde asparagine (N) située à l'extrémité C-terminale de la seconde hélice forme un liaison hydrogène avec la chaîne principale d'une isoleucine présente sur le brin  $\beta$ , stabilisant ainsi l'interaction entre ces deux structures secondaires (Groves *et al.*, 1996; Groves *et al.*, 1998).

Le motif «ERFNIN» peut cependant contenir certaines substitutions au niveau des résidus conservés. Par exemple, dans la cathepsine L, la phénylalanine est remplacée par un tryptophane. D'autres substitutions, retrouvées dans les propeptides des cathepsines F et W, sont caractérisées par le remplacement de l'isoleucine et l'asparagine, respectivement par une alanine et une glutamine. Ce motif, alors appelé «ERFNAQ», constitue une sous-classe appelée aussi «cathepsin F-like».

#### Le motif à trois tryptophanes

Dans les propeptides de la sous-famille L, en plus des interactions du motif «ERFNIN», l'assemblage des hélices  $\alpha$  1 et 2 est assuré par des interactions entre trois résidus aromatiques (Kreusch *et al.*, 2000) (Figure 11). Ces trois résidus localisés à l'intersection entre les deux hélices constituent le cœur hydrophobe du propeptide et sont entourés par d'autres chaînes latérales hydrophobes provenant des hélices  $\alpha$  1 et 2.

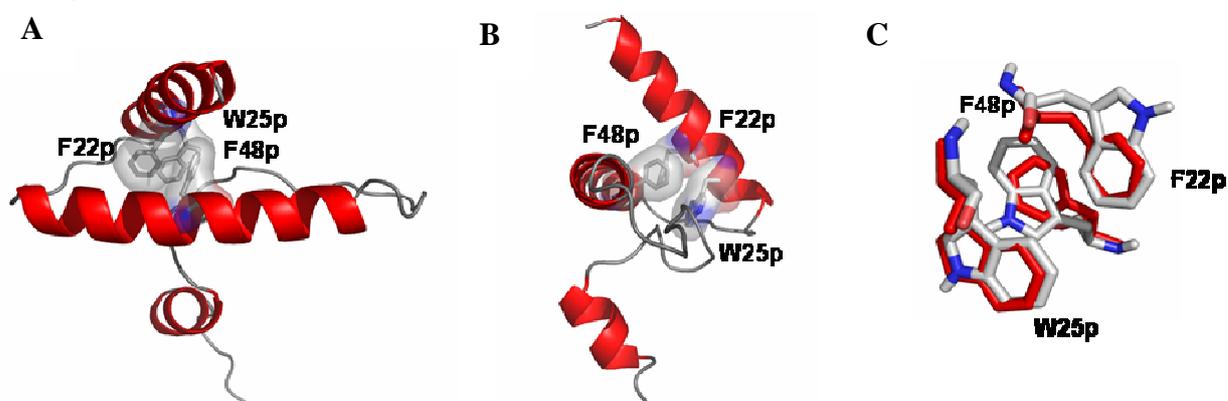


FIGURE 11: Localisation du motif à trois résidus aromatiques dans le propeptide de la caricaine. A et B: mise en évidence et localisation des différents résidus impliqués dans la formation du cœur aromatique du propeptide de la caricaine. Vue du dessus (A) et face latérale (B). C: comparaison de la localisation des trois résidus aromatiques F22p, W25p et F48p avec les trois tryptophanes du propeptide de la cathepsine L.

Dans le cœur aromatique du propeptide de la cathepsine L et S, le tryptophane (F/W) provenant de la seconde hélice et appartenant au motif «ER(F/W)NIN» est entouré des chaînes latérales de deux autres tryptophanes portés par l'hélice  $\alpha$ 1. Cette tripartite aromatique, conservée dans la sous-famille L, semble être très importante pour la stabilité du domaine N-terminal. En effet, Kreusch a montré que, chez la procathepsine S, la mutation de n'importe lequel de ces trois tryptophanes conduit à un mauvais repliement du zymogène ainsi qu'à une chute importante du potentiel inhibiteur du propeptide. Cependant, dans certains propeptides de la famille L-like, un ou plusieurs tryptophanes peuvent être substitués par d'autres résidus aromatiques, comme c'est le cas dans la caricaine où deux tryptophanes sont remplacés par des phénylalanines (Figure 11C).

#### Le motif GNFD

Un troisième élément structural, appelé motif «GNFD» est conservé dans la plupart des propeptides de la sous-famille L (Vernet *et al.*, 1995). Ce motif est localisé dans la boucle située entre le brin  $\beta$  et la troisième hélice. Ce motif, découvert par Vernet dans la papaïne, correspond à une alternance de résidus conservés et variables selon la séquence GXNXFXD.

Dans la papaïne, le motif GNFD jouerait un rôle important dans la modulation du processus d'activation car le remplacement de l'aspartate par une asparagine, une tyrosine, une méthionine, une

valine ou un glutamate conduit un zymogène non fonctionnel tandis que le remplacement de la phénylalanine par une histidine conduit à une augmentation d'une unité du pH optimum d'activation (Vernet *et al.*, 1995).

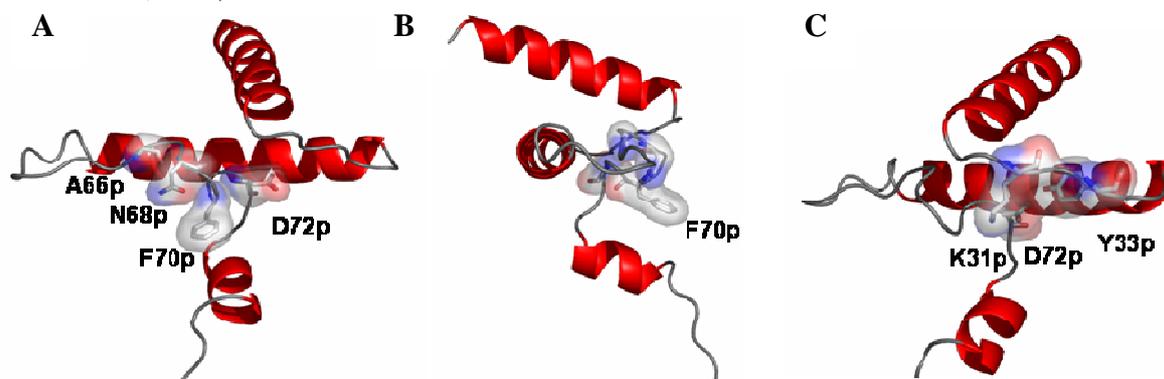


FIGURE 12: Localisation du motif GNFD dans le propeptide de la caricaine. A et B: mise en évidence et localisation des différents résidus impliqués dans le motif GNFD sur la face arrière (A) et latérale (B) du propeptide de la caricaine. C: mise en évidence de la triade Lys31p-Asp72p-Tyr33p reliant l'Asp72p situé dans la boucle entre les hélices  $\alpha$  2 et 3 et les résidus Kys31p et Tyr33p localisés dans la boucle située entre les hélices  $\alpha$  1 et 2.

Dans la structure du propeptide de la caricaine, l'aspartate du motif « GNFD » forme un pont salin avec une lysine et une liaison hydrogène avec une tyrosine (Figure 12C), tous deux présents dans la boucle située entre les hélices  $\alpha$  1 et 2 et pourrait donc stabiliser l'assemblage entre les trois hélices. Pour sa part, la chaîne latérale de la phénylalanine pointe vers la forme mature et est complètement enfouie dans la structure du zymogène. Dès lors, son remplacement par une histidine introduisant un groupe ionisable pourrait conduire à une déstabilisation de la surface d'interaction en dessous de pH 5 lorsque l'histidine est protonée (Groves *et al.*, 1996). A l'heure actuelle, la fonction de la glycine et de l'asparagine n'est pas connue. Cependant, une substitution de la glycine par une alanine est présente dans un grand nombre de propeptides tels que ceux de la cathepsin L et de caricaine.

#### Propeptide de la sous-famille B

##### Caractéristiques structurales

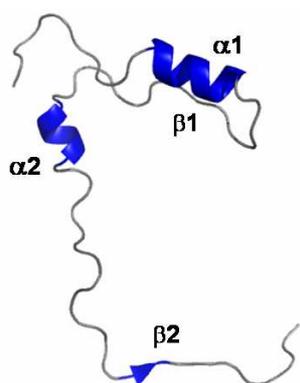


FIGURE 13: Structure du propeptide de la cathepsine B

Les propeptides de la famille B-like ne possèdent pas de motif ERFNIN. Ils sont significativement plus courts (60 acides aminés) que ceux de la sous-famille L ; leur repliement est différent et beaucoup moins compact (Groves *et al.*, 1998) (Figure 13). La partie N-terminale des propeptides B-like est formée d'une hélice  $\alpha$ , semblable à la deuxième hélice de la sous-famille L et d'un brin  $\beta$  antiparallèle qui forment ensemble une épingle à cheveux stabilisée par des interactions de type van der Waals. Les propeptides de la sous-famille B interagissent avec le domaine catalytique au niveau d'une région équivalente à la Propeptide Binding Loop (PBL) des protéases L-like.

Cependant, dans la sous-famille B, l'absence de l'hélice  $\alpha$ 1 est compensée par la présence d'une extension du domaine catalytique, appelée Occluding Loop. La partie C-terminale, tout comme celle des L-like, suit le sillon catalytique et masque le site catalytique.

### Similarité entre les sous-familles L et B

Bien que les propeptides des sous-familles L et B diffèrent significativement en taille et en homologie de séquence, leur repliement et les zones de la protéase mature avec lesquelles ils interagissent sont très similaires. En effet, les deux sous-familles possèdent un motif structural central en épingle à cheveux de type Helix-Turn-Strand qui est conservé malgré la faible homologie de séquences. Les interactions hydrophobe et hydrogène entre les deux faces de cette structure sont conservées entre les deux sous-familles. Une autre similarité entre les deux sous-familles est la présence d'un motif conservé de type two-turn- $\alpha$  helix situé en C-terminal du brin  $\beta$  et qui masque le sous-site S' de reconnaissance de substrat.

### Propeptide de Der p 1: une nouvelle sous-famille ?

#### Caractéristiques structurales

Le propeptide de Der p 1 est composé de 80 résidus et est donc caractérisé par une taille intermédiaire entre la taille des propeptides de la sous-famille B (60 résidus) et celle de la sous-famille L (100 résidus). La récente résolution de la structure tridimensionnelle du zymogène proDer p 1 par Meno a permis d'obtenir de nombreuses informations sur la structure du propeptide de Der p 1 et sur les interactions entre celui-ci et la protéase mature (Meno *et al.*, 2005). La structure montre que le propeptide de Der p 1 est formé de quatre hélices  $\alpha$  amphipatiques reliées par de courtes boucles, ce qui le distingue des propeptides de la sous-famille L composés de trois hélices et de la sous-famille B formée uniquement de deux hélices.

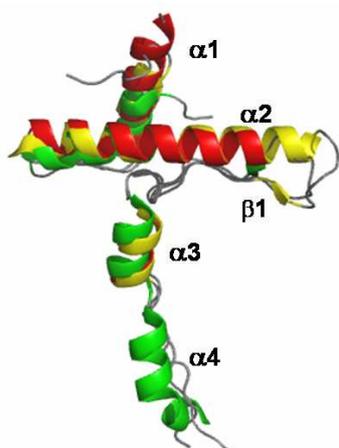


FIGURE 14: Alignement des structures des propeptides de la cathepsine L (jaune), de la caricaine (rouge) et de Der p 1 (vert)

Cependant, l'alignement de la structure du propeptide de Der p 1 avec les structures des propeptides de la cathepsine L humaine et de la caricaine, appartenant à la sous-famille «ERFNIN», montre que le propeptide de Der p 1 adopte un repliement global semblable à celui des propeptides de la sous-famille L.

Cet alignement (Figure 14) montre la conservation de la conformation adoptée par les hélices au niveau du domaine globulaire N-terminal malgré la faible homologie de séquence. De plus, les interactions aromatiques entre les deux résidus tryptophanes de la première hélice et celui du motif «ERFNIN», porté par la seconde hélice qui constitue le cœur hydrophobe et stabilise l'assemblage des hélices  $\alpha 1$  et  $\alpha 2$  du domaine N-terminal, sont conservées dans le propeptide de Der p 1 et correspondent aux résidus F8, Y11 et F32.

Néanmoins, plusieurs différences structurales distinguent le propeptide de Der p 1 des propeptides de la sous-famille L. En effet, le propeptide de Der p 1 ne contient pas de véritable motif «ERFNIN», il ne possède pas de brin  $\beta$  et présente une quatrième hélice à la place de la longue boucle C-terminale. L'alignement des structures montre que l'équivalent du motif  $EX_3RX_2(V/I)FX_2NX_3IX_3N$  dans le propeptide de Der p 1 serait la séquence **DEX<sub>2</sub>ARXNFX<sub>2</sub>SX<sub>3</sub>VX<sub>2</sub>NG** où les acides aminés en gras occupent les positions occupées par les résidus du motif ERFNIN dans la sous-famille L et ceux grisés correspondraient aux résidus attendus dans le véritable motif (Figure 15A).

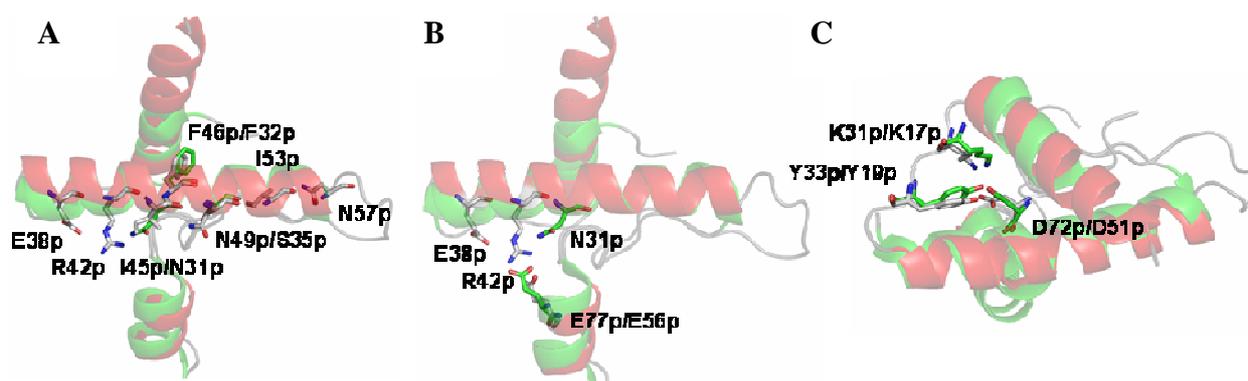


FIGURE 15: Comparaison des résidus conservés dans la stabilisation du domaine N-terminal du propeptide de la caricaine et de Der p 1. A: absence de conservation du motif ERFNIN dans le propeptide de Der p 1 (numérotation caricaine/Der p 1) B: absence de conservation de la triade Glu38p-Arg42p-Glu77p de la caricaine qui est remplacée dans le propeptide de Der p 1 par un pont hydrogène entre le Glu56p et l'Asn31p. C: conservation de la triade Lys31p-Asp72p-Tyr33p du propeptide de la caricaine dans le propeptide de Der p 1 où la triade correspondante implique les résidus Lys17p-Asp51p-Tyr19p.

Cette observation implique plusieurs différences entre le propeptide de Der p 1 et ceux de la famille L. Premièrement, alors que dans la sous-famille L, le domaine globulaire est stabilisé par des ponts salins et des liaisons hydrogène intra-propeptidiques entre les résidus E38p-R42p-E77p (Figure 15B) et K31p-D72p-Y33p (Figure 15C). Dans le propeptide de Der p 1, seules les interactions entre les résidus K17-D51-Y19 sont conservées, tandis que les résidus E38p et R42p, appartenant au motif ERFNIN, ne sont pas conservés. Deuxièmement, l'isoleucine ou la valine du motif ER(V/I)FNIN, qui fait partie du cœur hydrophobe du domaine globulaire N-terminal n'est pas conservée dans le propeptide de Der p 1 mais est remplacée par le résidu asparagine 31 dont la chaîne latérale pointe dans la direction opposée. Troisièmement, la première asparagine du motif ERFNIN, qui forme normalement une liaison hydrogène entre l'hélice  $\alpha_2$  et la chaîne principale de la boucle précédant l'hélice  $\alpha_3$ , n'est pas conservée dans le propeptide de Der p 1. Néanmoins, une interaction similaire pourrait être assurée par le résidu Sérine 35. Enfin, l'isoleucine et l'asparagine du motif ERFNIN, qui assurent normalement l'orientation relative de l'hélice  $\alpha_2$  par rapport au brin  $\beta$  par des interactions hydrophobe et des liaisons hydrogène, ne sont pas conservées. Cependant, cette absence pourrait être compensée par l'absence de brin  $\beta$  vraiment défini ainsi que par la taille plus courte de l'hélice  $\alpha_2$  du propeptide de Der p 1.

En conclusion, le seul résidu conservé du motif ERFNIN dans le propeptide de Der p 1 est le résidu phénylalanine en position 32. Cette observation ne permet pas de classer le propeptide de Der p 1 dans la sous-famille L. D'autre part, le propeptide de Der p 1 ne peut pas être classé dans la sous-famille B de par ses trop grandes différences structurales ainsi que par la présence de l'hélice  $\alpha_1$  et 4. Dès lors, le propeptide de Der p 1 ne peut être classé dans aucune de ces deux sous-familles. Dans ces conditions, Meno et ses collaborateurs ont proposé de le classer dans une troisième sous-famille qui serait caractérisée par un propeptide de taille intermédiaire, par la présence de quatre hélices  $\alpha$ , ainsi que par l'absence de motif « ERFNIN » et de brin  $\beta$  (Meno *et al.*, 2005).

## Mécanismes d'activation des zymogènes de la famille CA1

L'activation d'un zymogène de protéase correspond à la perte du propeptide et à l'apparition de l'activité protéolytique du domaine catalytique. Ce processus peut être catalysé par la molécule elle-même et on parle alors d'un mécanisme intramoléculaire (monomoléculaire) ou être catalysé par une autre molécule et, dans ce cas, sera appelé mécanisme intermoléculaire (bimoléculaire) (Rowan *et al.*, 1992).

### Activation *in vitro*

L'activation *in vitro* de zymogènes de protéases à cystéine active représente un déficit important dans l'obtention de ces protéases sous forme recombinantes car l'expression de la protéase ne peut être envisagée que sous la forme de son zymogène, dû au rôle de « chaperonne intramoléculaire » rempli par le propeptide. Il est donc nécessaire et obligatoire de déterminer et maîtriser les conditions dans lesquelles le processus d'activation peut être réalisé.

Un des premiers zymogènes recombinants de protéase à cystéine active à avoir été correctement mûré *in vitro* fut la propapaïne (Vernet *et al.*, 1991). Depuis, plusieurs études ont démontré qu'il était possible d'activer un grand nombre de zymogènes de protéases à cystéine active de la famille 1. Ces études, principalement réalisées avec des précurseurs de protéases lysosomales ont montré que la diminution du pH est essentielle pour obtenir une maturation. Celle-ci peut être ensuite réalisée par protéolyse intramoléculaire ou intermoléculaire selon le zymogène considéré (Bossard *et al.*, 1996; McQueney *et al.*, 1997; Quraishi and Storer, 2001). Ainsi, Menard *et al.* ont montré que l'activation de la procathepsine L humaine se déroule par un mécanisme inter et/ou intramoléculaire (Menard *et al.*, 1998). Dans ce cas, plusieurs sites de clivage, un situé à trois résidus de l'extrémité N-terminale et quatre situés à proximité de l'extrémité C-terminale ont pu être identifiés. Ces sites correspondent tous à la spécificité protéolytique de la cathepsine L mature et leur localisation sur la structure tridimensionnelle du précurseur montre qu'ils sont tous situés dans des zones exposées au solvant. Des résultats équivalents ont été obtenus avec la procathepsine K (McQueney *et al.*, 1997). D'autre part, Quraishi *et al.* en bloquant l'activation de la procathepsine B et S par de la cystatine, ont identifié deux sites de clivages internes, équivalents dans les deux précurseurs, situés au niveau de la zone d'interaction avec le site catalytique (Quraishi and Storer, 2001). La présence de cystatine et la localisation de ces sites sembleraient indiquer qu'ils pourraient provenir d'un processus intramoléculaire.

*In vitro*, ces processus d'activation peuvent parfois conduire à la formation de forme mature avec une extension de quelques acides aminés comme c'est le cas pour la cathepsine K et S et la papaïne, suggérant que d'autres protéases pourraient être nécessaires pour terminer le processus de maturation (Vernet *et al.*, 1990; Linnevers *et al.*, 1997). Enfin, certains zymogènes nécessitent obligatoirement l'intervention d'une autre protéase pour être transformés en protéases mature comme c'est le cas de la PPIV. Pour cette dernière, l'absence d'autoactivation proviendrait de la trop grande spécificité du domaine catalytique, ne lui permettant pas de reconnaître et de cliver son propeptide (Baker *et al.*, 1996).

### Activation *in vivo*

Les conditions d'activation *in vivo* d'un zymogène de protéase à cystéine active peuvent être très complexes et peuvent requérir outre une diminution du pH, la présence d'autres protéases et/ou l'interaction avec des macromolécules comme le glucosaminoglycan (Ishidoh and Kominami, 1995; Kihara *et al.*, 2002). Ce processus est réalisé presque exclusivement à l'intérieur de la cellule dans les lysosomes ou dans des endosomes bien que des activations extracellulaires puissent avoir lieu (Wiederanders, 2003). Par exemple, la procathepsine B humaine peut s'autoactiver *in vivo* (Mach *et al.*, 1993) mais aussi être activée par le mélange complexe des cathepsines H, S, C et L présentes dans le lysosome (Ishidoh *et al.*, 1999). D'autre part, la procathepsine B peut aussi être activée, après sa sécrétion, dans la matrice extracellulaire par l'action de la cathepsine D (van der Stappen *et al.*, 1996) illustrant la complexité et la diversité des processus d'activation qui peuvent intervenir.

### Mécanisme d'activation du zymogène proDer p 1

Durant la dernière décennie, plusieurs essais ont été réalisés afin de déterminer les conditions d'activation du zymogène proDer p 1 en protéase mature. Cette étape est, en effet, essentielle pour obtenir une protéase Der p 1 recombinante active et présentant une reconnaissance par les IgE équivalente à celle de l'allergène naturel. Ces essais ont été réalisés, avec plus ou moins de succès, à partir de précurseurs exprimés dans différents organismes en appliquant des conditions d'activation préalablement décrites pour l'activation d'autres protéases à cystéine active de la famille CA1. Ces différentes activations, bien que réalisées dans des conditions différentes, ont permis d'obtenir des informations sur le processus d'activation.

La première activation du précurseur de Der p 1 a été décrite par Jacquet *et al.* (Jacquet *et al.*, 2002) après un traitement à pH acide d'un zymogène proDer p 1 exprimé par la levure *Pichia pastoris*. Cette activation fut réalisée à partir d'un zymogène hyperglycosylé et partiellement activé dans lequel les 18 premiers résidus (ATFE-) correspondant à la première hélice du propeptide étaient délétés au terme de la production. D'autre part, la protéase générée après activation se présentait sous la forme d'un mélange de protéases avec et sans extension de deux résidus (AETNA-/TNA-). En complément de cette étude, la comparaison de la reconnaissance par les IgE du précurseur proDer p 1 avec la protéase mature Der p 1 et l'allergène naturel a permis de mettre en évidence le caractère « hypoallergénique » du zymogène vis-à-vis de la protéase mature. Cette observation fut confirmée par d'autres groupes, suggérant que des épitopes majeurs de la protéase puissent être masqués par le propeptide dans le précurseur. Par la suite, la comparaison de l'activation à pH acide du zymogène proDer p 1 avec celle du zymogène proDer p 1 possédant un site potentiel de N-glycosylation supplémentaire dans le propeptide mit en évidence, pour ces deux précurseurs, la présence d'un intermédiaire correspondant à la perte de la première hélice du propeptide. Cette étude suggéra que la glycosylation du propeptide puisse inhiber le phénomène d'activation du zymogène proDer p 1 (Takai *et al.*, 2002). Cette hypothèse fut renforcée après l'obtention de Der p 1 mature après déglycosylation enzymatique à pH 5,5 d'un zymogène exprimé par la levure *Pichia pastoris* ou après incubation avec l'allergène naturel immobilisé sur un support solide suggérant que des phénomènes de protéolyse intermoléculaire puissent avoir lieu lors du processus d'activation (van Oort *et al.*, 2002).



- Anfinsen, C. B.** (1973). Principles that govern the folding of protein chains. *Science* 181(96): 223-30.
- Baker, K. C., M. A. Taylor, N. J. Cummings, M. A. Tunon, K. A. Worboys and I. F. Connerton** (1996). Autocatalytic processing of pro-papaya proteinase IV is prevented by crowding of the active-site cleft. *Protein Eng* 9(6): 525-9.
- Baranski, T. J., A. B. Cantor and S. Kornfeld** (1992). Lysosomal enzyme phosphorylation. I. Protein recognition determinants in both lobes of procathepsin D mediate its interaction with UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase. *J Biol Chem* 267(32): 23342-8.
- Baranski, T. J., G. Koelsch, J. A. Hartsuck and S. Kornfeld** (1991). Mapping and molecular modeling of a recognition domain for lysosomal enzyme targeting. *J Biol Chem* 266(34): 23365-72.
- Barrett, A., Rawlings and Woessner** (1998). Cysteine peptidases. In *Handbook of Proteolytic Enzymes*. Academic Press, London: 543-798.
- Bennett, B. J. and W. R. Thomas** (1996). Cloning and sequencing of the group 6 allergen of *Dermatophagoides pteronyssinus*. *Clin Exp Allergy* 26(10): 1150-4.
- Best, E. A., K. E. Stedman, C. M. Bozic, S. W. Hunter, L. Vailes, M. D. Chapman, C. A. McCall and M. J. McDermott** (2000). A recombinant group 1 house dust mite allergen, rDer f 1, with biological activities similar to those of the native allergen. *Protein Expr Purif* 20(3): 462-71.
- Billington, C. J., P. Mason, M. C. Magny and J. S. Mort** (2000). The slow-binding inhibition of cathepsin K by its propeptide. *Biochem Biophys Res Commun* 276(3): 924-9.
- Bossard, M. J., T. A. Tomaszek, S. K. Thompson, B. Y. Amegadzie, C. R. Hanning, C. Jones, J. T. Kurdyla, D. E. McNulty, F. H. Drake, M. Gowen and M. A. Levy** (1996). Proteolytic activity of human osteoclast cathepsin K. Expression, purification, activation, and substrate identification. *J Biol Chem* 271(21): 12517-24.
- Brown, A., K. Farmer, L. MacDonald, N. Kalsheker, D. Pritchard, C. Haslett, J. Lamb and J. M. Sallenave** (2003). House dust mite Der p 1 downregulates defenses of the lung by inactivating elastase inhibitors. *Am J Respir Cell Mol Biol* 29(3 Pt 1): 381-9.
- Carmona, E., E. Dufour, C. Plouffe, S. Takebe, P. Mason, J. S. Mort and R. Menard** (1996). Potency and selectivity of the cathepsin L propeptide as an inhibitor of cysteine proteases. *Biochemistry* 35(25): 8149-57.
- Chapman, H. A., R. J. Riese and G. P. Shi** (1997). Emerging roles for cysteine proteases in human biology. *Annu Rev Physiol* 59: 63-88.
- Chapman, M. D. and T. A. Platts-Mills** (1980). Purification and characterization of the major allergen from *Dermatophagoides pteronyssinus*-antigen P1. *J Immunol* 125(2): 587-92.
- Chapman, M. D., S. Rowntree, E. B. Mitchell, M. C. Di Prisco de Fuenmajor and T. A. Platts-Mills** (1983). Quantitative assessments of IgG and IgE antibodies to inhalant allergens in patients with atopic dermatitis. *J Allergy Clin Immunol* 72(1): 27-33.
- Choi, K. H., R. A. Laursen and K. N. Allen** (1999). The 2.1 Å structure of a cysteine protease with proline specificity from ginger rhizome, *Zingiber officinale*. *Biochemistry* 38(36): 11624-33.
- Chua, K. Y., P. K. Kehal, W. R. Thomas, P. R. Vaughan and I. G. Macreadie** (1992). High-frequency binding of IgE to the Der p allergen expressed in yeast. *J Allergy Clin Immunol* 89(1 Pt 1): 95-102.
- Chua, K. Y., G. A. Stewart, W. R. Thomas, R. J. Simpson, R. J. Dilworth, T. M. Plozza and K. J. Turner** (1988). Sequence analysis of cDNA coding for a major house dust mite allergen, Der p 1. Homology with cysteine proteases. *J Exp Med* 167(1): 175-82.
- Comoy, E. E., J. Pestel, C. Duez, G. A. Stewart, C. Vendeville, C. Fournier, F. Finkelman, A. Capron and G. Thyphronitis** (1998). The house dust mite allergen, *Dermatophagoides pteronyssinus*, promotes type 2 responses by modulating the balance between IL-4 and IFN-gamma. *J Immunol* 160(5): 2456-62.
- Cookson, W., M. Moffatt and** (2000). Genetics of asthma and allergic diseases. *Hum Mol Gen.* 9: 2359-2364.
- Cookson, W. O., P. A. Sharp, J. A. Faux and J. M. Hopkin** (1989). Linkage between immunoglobulin E responses underlying asthma and rhinitis and chromosome 11q. *Lancet* 1(8650): 1292-5.
- Coombs, R. and P. Gell** (1975). Classification of allergic reactions responsible for clinical hypersensitivity and disease. Oxford: Blackwell Scientific.
- Coulombe, R., P. Grochulski, J. Sivaraman, R. Menard, J. S. Mort and M. Cygler** (1996). Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *Embo J* 15(20): 5492-503.
- Cuozzo, J. W., K. Tao, M. Cygler, J. S. Mort and G. G. Sahagian** (1998). Lysine-based structure responsible for selective mannose phosphorylation of cathepsin D and cathepsin L defines a common structural motif for lysosomal enzyme targeting. *J Biol Chem* 273(33): 21067-76.

- Cuozzo, J. W., K. Tao, Q. L. Wu, W. Young and G. G. Sahagian** (1995). Lysine-based structure in the proregion of procathepsin L is the recognition site for mannose phosphorylation. *J Biol Chem* 270(26): 15611-9.
- Cyglar, M., J. Sivaraman, P. Grochulski, R. Coulombe, A. C. Storer and J. S. Mort** (1996). Structure of rat procathepsin B: model for inhibition of cysteine protease activity by the proregion. *Structure* 4(4): 405-16.
- de Halleux, S., E. Stura, L. VanderElst, V. Carlier, M. Jacquemin and J. M. Saint-Remy** (2006). Three-dimensional structure and IgE-binding properties of mature fully active Der p 1, a clinically relevant major allergen. *J Allergy Clin Immunol* 117(3): 571-6.
- Deb, R., F. Shakib, K. Reid and H. Clark** (2007). Major house dust mite allergens *Dermatophagoides pteronyssinus* 1 and *Dermatophagoides farinae* 1 degrade and inactivate lung surfactant proteins A and D. *J Biol Chem* 282(51): 36808-19.
- Del Prete, G., E. Maggi, P. Parronchi, I. Chretien, A. Tiri, D. Macchia, M. Ricci, J. Banchereau, J. De Vries and S. Romagnani** (1988). IL-4 is an essential factor for the IgE synthesis induced in vitro by human T cell clones and their supernatants. *J Immunol* 140(12): 4193-8.
- Deslee, G., A. S. Charbonnier, H. Hammad, G. Angyalosi, I. Tillie-Leblond, A. Mantovani, A. B. Tonnel and J. Pestel** (2002). Involvement of the mannose receptor in the uptake of Der p 1, a major mite allergen, by human dendritic cells. *J Allergy Clin Immunol* 110(5): 763-70.
- Dickinson, D. P.** (2002). Cysteine peptidases of mammals: their biological roles and potential effects in the oral cavity and other tissues in health and disease. *Crit Rev Oral Biol Med* 13(3): 238-75.
- Dilworth, R. J., K. Y. Chua and W. R. Thomas** (1991). Sequence analysis of cDNA coding for a major house dust mite allergen, Der f I. *Clin Exp Allergy* 21(1): 25-32.
- Drenth, J., J. N. Jansonius, R. Koekoek, H. M. Swen and B. G. Wolthers** (1968). Structure of papain. *Nature* 218(5145): 929-32.
- Fox, T., E. de Miguel, J. S. Mort and A. C. Storer** (1992). Potent slow-binding inhibition of cathepsin B by its propeptide. *Biochemistry* 31(50): 12571-6.
- Fujishima, A., Y. Imai, T. Nomura, Y. Fujisawa, Y. Yamamoto and T. Sugawara** (1997). The crystal structure of human cathepsin L complexed with E-64. *FEBS Lett* 407(1): 47-50.
- Furmonaviciene, R., A. M. Ghaemmaghani, S. E. Boyd, N. S. Jones, K. Bailey, A. C. Willis, H. F. Sewell, D. A. Mitchell and F. Shakib** (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. *Clin Exp Allergy* 37(2): 231-42.
- Ghaemmaghani, A. M. and F. Shakib** (2002). Human T cells that have been conditioned by the proteolytic activity of the major dust mite allergen Der p 1 trigger enhanced immunoglobulin E synthesis by B cells. *Clin Exp Allergy* 32(5): 728-32.
- Goldsby, R., T. Kindt and B. Osborne** (2001). *Immunologie, le cours de Janis Kuby*. Paris, Freeman and compagny.
- Gough, L., O. Schulz, H. F. Sewell and F. Shakib** (1999). The cysteine protease activity of the major dust mite allergen Der p 1 selectively enhances the immunoglobulin E antibody response. *J Exp Med* 190(12): 1897-902.
- Groves, M. R., R. Coulombe, J. Jenkins and M. Cyglar** (1998). Structural basis for specificity of papain-like cysteine protease proregions toward their cognate enzymes. *Proteins* 32(4): 504-14.
- Groves, M. R., M. A. Taylor, M. Scott, N. J. Cummings, R. W. Pickersgill and J. A. Jenkins** (1996). The prosequence of procaricain forms an alpha-helical domain that prevents access to the substrate-binding cleft. *Structure* 4(10): 1193-203.
- Guay, J., J. P. Falguyret, A. Ducret, M. D. Percival and J. A. Mancini** (2000). Potency and selectivity of inhibition of cathepsin K, L and S by their respective propeptides. *Eur J Biochem* 267(20): 6311-8.
- Harris, J., D. E. Mason, J. Li, K. W. Burdick, B. J. Backes, T. Chen, A. Shipway, G. Van Heeke, L. Gough, A. Ghaemmaghani, F. Shakib, F. Debaene and N. Winssinger** (2004). Activity profile of dust mite allergen extract using substrate libraries and functional proteomic microarrays. *Chem Biol* 11(10): 1361-72.
- Huete-Perez, J. A., J. C. Engel, L. S. Brinen, J. C. Mottram and J. H. McKerrow** (1999). Protease trafficking in two primitive eukaryotes is mediated by a prodomain protein motif. *J Biol Chem* 274(23): 16249-56.
- Ishidoh, K. and E. Kominami** (1995). Procathepsin L degrades extracellular matrix proteins in the presence of glycosaminoglycans in vitro. *Biochem Biophys Res Commun* 217(2): 624-31.
- Ishidoh, K., M. Takeda-Ezaki, S. Watanabe, N. Sato, M. Aihara, K. Imagawa, M. Kikuchi and E. Kominami** (1999). Analysis of where and which types of proteinases participate in lysosomal proteinase processing using bafilomycin A1 and *Helicobacter pylori* Vac A toxin. *J Biochem* 125(4): 770-9.

- Jacquet, A., M. Haumont, M. Massaer, V. Daminet, L. Garcia, P. Mazzu, P. Jacobs and A. Bollen** (2000). Biochemical and immunological characterization of a recombinant precursor form of the house dust mite allergen Der p 1 produced by *Drosophila* cells. *Clin Exp Allergy* 30(5): 677-84.
- Jacquet, A., M. Magi, H. Petry and A. Bollen** (2002). High-level expression of recombinant house dust mite allergen Der p 1 in *Pichia pastoris*. *Clin Exp Allergy* 32(7): 1048-53.
- Janeway, C., Travers P, M. Walport and J. Capra** (1999). *The Immunesystem in Health and Disease*. New York/London., Garland publishing Inc.
- Kalsheker, N. A., S. Deam, L. Chambers, S. Sreedharan, K. Brocklehurst and D. A. Lomas** (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. *Biochem Biophys Res Commun* 221(1): 59-61.
- Karrer, K. M., S. L. Peiffer and M. E. DiTomas** (1993). Two distinct gene subfamilies within the family of cysteine protease genes. *Proc Natl Acad Sci U S A* 90(7): 3063-7.
- Katerelos, N. A., M. A. Taylor, M. Scott, P. W. Goodenough and R. W. Pickersgill** (1996). Crystal structure of a caricain D158E mutant in complex with E-64. *FEBS Lett* 392(1): 35-9.
- Khan, A. R. and M. N. James** (1998). Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Sci* 7(4): 815-36.
- Kihara, M., H. Kakegawa, Y. Matano, E. Murata, H. Tsuge, H. Kido and N. Katunuma** (2002). Chondroitin sulfate proteoglycan is a potent enhancer in the processing of procathepsin L. *Biol Chem* 383(12): 1925-9.
- Kikuchi, Y., T. Takai, T. Kuhara, M. Ota, T. Kato, H. Hatanaka, S. Ichikawa, T. Tokura, H. Akiba, K. Mitsuishi, S. Ikeda, K. Okumura and H. Ogawa** (2006). Crucial commitment of proteolytic activity of a purified recombinant major house dust mite allergen Der p1 to sensitization toward IgE and IgG responses. *J Immunol* 177(3): 1609-17.
- King, C., R. J. Simpson, R. L. Moritz, G. E. Reed, P. J. Thompson and G. A. Stewart** (1996). The isolation and characterization of a novel collagenolytic serine protease allergen (Der p 9) from the dust mite *Dermatophagoides pteronyssinus*. *J Allergy Clin Immunol* 98(4): 739-47.
- Kreusch, S., M. Fehn, G. Maubach, K. Nissler, W. Rommerskirch, K. Schilling, E. Weber, I. Wenz and B. Wiederanders** (2000). An evolutionarily conserved tripartite tryptophan motif stabilizes the prodomains of cathepsin L-like cysteine proteases. *Eur J Biochem* 267(10): 2965-72.
- Krillis, S., B. A. Baldo, R. Sutton and A. Basten** (1984). Antigens and allergens from the common house dust mite *Dermatophagoides pteronyssinus*. Part I. Demonstration of multiple allergens by immunochemical and biologic analyses. *J Allergy Clin Immunol* 74(2): 132-41.
- Lind, P.** (1985). Purification and partial characterization of two major allergens from the house dust mite *Dermatophagoides pteronyssinus*. *J Allergy Clin Immunol* 76(5): 753-61.
- Linnevers, C. J., M. E. McGrath, R. Armstrong, F. R. Mistry, M. G. Barnes, J. L. Klaus, J. T. Palmer, B. A. Katz and D. Bromme** (1997). Expression of human cathepsin K in *Pichia pastoris* and preliminary crystallographic studies of an inhibitor complex. *Protein Sci* 6(4): 919-21.
- Mach, L., H. Schwihla, K. Stuwe, A. D. Rowan, J. S. Mort and J. Glossl** (1993). Activation of procathepsin B in human hepatoma cells: the conversion into the mature enzyme relies on the action of cathepsin B itself. *Biochem J* 293 ( Pt 2): 437-42.
- Maruo, K., T. Akaike, T. Ono, T. Okamoto and H. Maeda** (1997). Generation of anaphylatoxins through proteolytic processing of C3 and C5 by house dust mite protease. *J Allergy Clin Immunol* 100(2): 253-60.
- Massaer, M., P. Mazzu, M. Haumont, M. Magi, V. Daminet, A. Bollen and A. Jacquet** (2001). High-level expression in mammalian cells of recombinant house dust mite allergen ProDer p 1 with optimized codon usage. *Int Arch Allergy Immunol* 125(1): 32-43.
- Maubach, G., K. Schilling, W. Rommerskirch, I. Wenz, J. E. Schultz, E. Weber and B. Wiederanders** (1997). The inhibition of cathepsin S by its propeptide--specificity and mechanism of action. *Eur J Biochem* 250(3): 745-50.
- McGrath, M. E., J. L. Klaus, M. G. Barnes and D. Bromme** (1997). Crystal structure of human cathepsin K complexed with a potent inhibitor. *Nat Struct Biol* 4(2): 105-9.
- McIntyre, G. F. and A. H. Erickson** (1993). The lysosomal proenzyme receptor that binds procathepsin L to microsomal membranes at pH 5 is a 43-kDa integral membrane protein. *Proc Natl Acad Sci U S A* 90(22): 10588-92.
- McIntyre, G. F., G. D. Godbold and A. H. Erickson** (1994). The pH-dependent membrane association of procathepsin L is mediated by a 9-residue sequence within the propeptide. *J Biol Chem* 269(1): 567-72.

- McQueney, M. S., B. Y. Amegadzie, K. D'Alessio, C. R. Hanning, M. M. McLaughlin, D. McNulty, S. A. Carr, C. Ijames, J. Kurdyla and C. S. Jones** (1997). Autocatalytic activation of human cathepsin K. *J Biol Chem* 272(21): 13955-60.
- Mehtani, S., Q. Gong, J. Panella, S. Subbiah, D. M. Peffley and A. Frankfater** (1998). In vivo expression of an alternatively spliced human tumor message that encodes a truncated form of cathepsin B. Subcellular distribution of the truncated enzyme in COS cells. *J Biol Chem* 273(21): 13236-44.
- Menard, R., E. Carmona, S. Takebe, E. Dufour, C. Plouffe, P. Mason and J. S. Mort** (1998). Autocatalytic processing of recombinant human procathepsin L. Contribution of both intermolecular and unimolecular events in the processing of procathepsin L in vitro. *J Biol Chem* 273(8): 4478-84.
- Meno, K., P. B. Thorsted, H. Ipsen, O. Kristensen, J. N. Larsen, M. D. Spangfort, M. Gajhede and K. Lund** (2005). The crystal structure of recombinant proDer p 1, a major house dust mite proteolytic allergen. *J Immunol* 175(6): 3835-45.
- Meno, K., P. B. Thorsted, H. Ipsen, O. Kristensen, J. N. Larsen, M. D. Spangfort, M. Gajhede and K. Lund** (2006). The metal binding site of the major house dust mite allergen Der p 1. *J Allergy Clin Immunol* 118(4): 971.
- Metz-Fabre, Rame J.-M., de Blay F. and P. G** (2006). Intérêt des allergènes recombinants pour la prise en charge de patients allergiques : cas cliniques. *Revue française d'allergologie et d'immunologie clinique*(46): 170-173.
- Meyer, C. H., J. F. Bond, M. S. Chen and M. T. Kasaian** (1994). Comparison of the levels of the major allergens Der p I and Der p II in standardized extracts of the house dust mite, *Dermatophagoides pteronyssinus*. *Clin Exp Allergy* 24(11): 1041-8.
- Miescher, S. M. and M. Vogel** (2002). Molecular aspects of allergy. *Mol Aspects Med* 23(6): 413-62.
- Miyamoto, T., S. Oshima, T. Ishizaki and S. H. Sato** (1968). Allergenic identity between the common floor mite (*Dermatophagoides farinae* Hughes, 1961) and house dust as a causative antigen in bronchial asthma. *J Allergy* 42(1): 14-28.
- Musil, D., D. Zucic, D. Turk, R. A. Engh, I. Mayr, R. Huber, T. Popovic, V. Turk, T. Towatari, N. Katunuma and et al.** (1991). The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *Embo J* 10(9): 2321-30.
- Mygind, N., R. Dahl, S. Pedersens and K. Thestrup-Pedersen** (1996). *Essential Allergy*. Blackwell Science Ltd.
- Nathanson, M. E.** (1969). Arthropods and allergy. *J Asthma Res* 7(2): 55-63.
- Neurath, H.** (1984). Evolution of proteolytic enzymes. *Science* 224(4647): 350-7.
- Ogino, T., T. Kaji, M. Kawabata, K. Satoh, K. Tomoo, T. Ishida, H. Yamazaki, K. Ishidoh and E. Kominami** (1999). Function of the propeptide region in recombinant expression of active procathepsin L in *Escherichia coli*. *J Biochem* 126(1): 78-83.
- Pene, J., F. Rousset, F. Briere, I. Chretien, J. Y. Bonnefoy, H. Spits, T. Yokota, N. Arai, K. Arai, J. Banchereau and et al.** (1988). IgE production by normal human lymphocytes is induced by interleukin 4 and suppressed by interferons gamma and alpha and prostaglandin E2. *Proc Natl Acad Sci U S A* 85(18): 6880-4.
- Peters, R. J., A. K. Shiau, J. L. Sohl, D. E. Anderson, G. Tang, J. L. Silen and D. A. Agard** (1998). Pro region C-terminus:protease active site interactions are critical in catalyzing the folding of alpha-lytic protease. *Biochemistry* 37(35): 12058-67.
- Piboonpocanun, S., N. Malainual, O. Jirapongsananuruk, P. Vichyanond and W. R. Thomas** (2006). Genetic polymorphisms of major house dust mite allergens. *Clin Exp Allergy* 36(4): 510-6.
- Pietschmann, S., M. Fehn, G. Kaulmann, I. Wenz, B. Wiederanders and K. Schilling** (2002). Foldase function of the cathepsin S proregion is strictly based upon its domain structure. *Biol Chem* 383(9): 1453-8.
- Quraishi, O. and A. C. Storer** (2001). Identification of internal autoproteolytic cleavage sites within the prosegments of recombinant procathepsin B and procathepsin S. Contribution of a plausible unimolecular autoproteolytic event for the processing of zymogens belonging to the papain family. *J Biol Chem* 276(11): 8118-24.
- Rawlings, N. D. and A. J. Barrett** (2000). MEROPS: the peptidase database. *Nucleic Acids Res* 28(1): 323-5.
- Reddy, V. Y., Q. Y. Zhang and S. J. Weiss** (1995). Pericellular mobilization of the tissue-destructive cysteine proteinases, cathepsins B, L, and S, by human monocyte-derived macrophages. *Proc Natl Acad Sci U S A* 92(9): 3849-53.
- Rowan, A. D., P. Mason, L. Mach and J. S. Mort** (1992). Rat procathepsin B. Proteolytic processing to the mature form in vitro. *J Biol Chem* 267(22): 15993-9.
- Ruffilli, A. and S. Bonini** (1997). Susceptibility genes for allergy and asthma. *Allergy* 52(3): 256-73.

- Sajid, M. and J. H. McKerrow** (2002). Cysteine proteases of parasitic organisms. *Mol Biochem Parasitol* 120(1): 1-21.
- Schechter, I. and A. Berger** (1967). On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* 27(2): 157-62.
- Schlereth, A., D. Standhardt, H. P. Mock and K. Muntz** (2001). Stored cysteine proteinases start globulin mobilization in protein bodies of embryonic axes and cotyledons during vetch (*Vicia sativa* L.) seed germination. *Planta* 212(5-6): 718-27.
- Schulz, O., P. Laing, H. F. Sewell and F. Shakib** (1995). Der p I, a major allergen of the house dust mite, proteolytically cleaves the low-affinity receptor for human IgE (CD23). *Eur J Immunol* 25(11): 3191-4.
- Schulz, O., H. F. Sewell and F. Shakib** (1998a). Proteolytic cleavage of CD25, the alpha subunit of the human T cell interleukin 2 receptor, by Der p 1, a major mite allergen with cysteine protease activity. *J Exp Med* 187(2): 271-5.
- Schulz, O., H. F. Sewell and F. Shakib** (1998b). A sensitive fluorescent assay for measuring the cysteine protease activity of Der p 1, a major allergen from the dust mite *Dermatophagoides pteronyssinus*. *Mol Pathol* 51(4): 222-4.
- Schulz, O., B. J. Sutton, R. L. Beavil, J. Shi, H. F. Sewell, H. J. Gould, P. Laing and F. Shakib** (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. *Eur J Immunol* 27(3): 584-8.
- Shakib, F., O. Schulz and H. Sewell** (1998). A mite subversive: cleavage of CD23 and CD25 by Der p 1 enhances allergenicity. *Immunol Today* 19(7): 313-6.
- Sharma, S., P. M. Lackie and S. T. Holgate** (2003). Uneasy breather: the implications of dust mite allergens. *Clin Exp Allergy* 33(2): 163-5.
- Shinde, U., Y. Li, S. Chatterjee and M. Inouye** (1993). Folding pathway mediated by an intramolecular chaperone. *Proc Natl Acad Sci U S A* 90(15): 6924-8.
- Sivaraman, J., M. Lalumiere, R. Menard and M. Cygler** (1999). Crystal structure of wild-type human procathepsin K. *Protein Sci* 8(2): 283-90.
- Smith, S. M. and M. M. Gottesman** (1989). Activity and deletion analysis of recombinant human cathepsin L expressed in *Escherichia coli*. *J Biol Chem* 264(34): 20487-95.
- Smith, W. A., K. Y. Chua, M. C. Kuo, B. L. Rogers and W. R. Thomas** (1994). Cloning and sequencing of the *Dermatophagoides pteronyssinus* group III allergen, Der p III. *Clin Exp Allergy* 24(3): 220-8.
- Smith, W. A., B. J. Hales, A. G. Jarnicki and W. R. Thomas** (2001). Allergens of wild house dust mites: environmental Der p 1 and Der p 2 sequence polymorphisms. *J Allergy Clin Immunol* 107(6): 985-92.
- Stewart, G. A., L. D. Ward, R. J. Simpson and P. J. Thompson** (1992). The group III allergen from the house dust mite *Dermatophagoides pteronyssinus* is a trypsin-like enzyme. *Immunology* 75(1): 29-35.
- Sun, G., M. A. Stacey, M. Schmidt, L. Mori and S. Mattoli** (2001). Interaction of mite allergens Der p3 and Der p9 with protease-activated receptor-2 expressed by lung epithelial cells. *J Immunol* 167(2): 1014-21.
- Takahashi, K., T. Takai, T. Yasuhara, T. Yuuki, Y. Ohtake, T. Yokota and Y. Okumura** (2000). Production of enzymatically and immunologically active Der f 1 in *Escherichia coli*. *Int Arch Allergy Immunol* 122(2): 108-14.
- Takai, T., T. Kato, Y. Sakata, H. Yasueda, K. Izuhara, K. Okumura and H. Ogawa** (2005). Recombinant Der p 1 and Der f 1 exhibit cysteine protease activity but no serine protease activity. *Biochem Biophys Res Commun* 328(4): 944-52.
- Takai, T., R. Mineki, T. Nakazawa, M. Takaoka, H. Yasueda, K. Murayama, K. Okumura and H. Ogawa** (2002). Maturation of the activities of recombinant mite allergens Der p 1 and Der f 1, and its implication in the blockade of proteolytic activity. *FEBS Lett* 531(2): 265-72.
- Tao, K., N. A. Stearns, J. Dong, Q. L. Wu and G. G. Sahagian** (1994). The proregion of cathepsin L is required for proper folding, stability, and ER exit. *Arch Biochem Biophys* 311(1): 19-27.
- Taylor, M. A., K. C. Baker, G. S. Briggs, I. F. Connerton, N. J. Cummings, K. A. Pratt, D. F. Revell, R. B. Freedman and P. W. Goodenough** (1995). Recombinant pro-regions from papain and papaya proteinase IV-are selective high affinity inhibitors of the mature papaya enzymes. *Protein Eng* 8(1): 59-62.
- Thomas, B., P. Heap and F. Carswell** (1991). Ultrastructural localization of the allergen Der p I in the gut of the house dust mite *Dermatophagoides pteronyssinus*. *Int Arch Allergy Appl Immunol* 94(1-4): 365-7.
- Thomas, W. R. and W. Smith** (1998). House-dust-mite allergens. *Allergy* 53(9): 821-32.
- Thomas, W. R., W. A. Smith, B. J. Hales, K. L. Mills and R. M. O'Brien** (2002). Characterization and immunobiology of house dust mite allergens. *Int Arch Allergy Immunol* 129(1): 1-18.

- Thomas, W. R., G. A. Stewart, R. J. Simpson, K. Y. Chua, T. M. Plozza, R. J. Dilworth, A. Nisbet and K. J. Turner** (1988). Cloning and expression of DNA coding for the major house dust mite allergen Der p 1 in *Escherichia coli*. *Int Arch Allergy Appl Immunol* 85(1): 127-9.
- Tovey, E. R., M. D. Chapman and T. A. Platts-Mills** (1981). Mite faeces are a major source of house dust allergens. *Nature* 289(5798): 592-3.
- Turk, B., V. Turk and D. Turk** (1997). Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol Chem* 378(3-4): 141-50.
- Twining, S. S.** (1994). Regulation of proteolytic activity in tissues. *Crit Rev Biochem Mol Biol* 29(5): 315-83.
- Valenta, R.** (2002). Recombinant allergen-based concepts for diagnosis and therapy of type I allergy. *Allergy* 57 Suppl 71: 66-7.
- van der Stappen, J. W., A. C. Williams, R. A. Maciewicz and C. Paraskeva** (1996). Activation of cathepsin B, secreted by a colorectal cancer cell line requires low pH and is mediated by cathepsin D. *Int J Cancer* 67(4): 547-54.
- van der Zee, J. S., H. de Groot, P. van Swieten, H. M. Jansen and R. C. Aalberse** (1988). Discrepancies between the skin test and IgE antibody assays: study of histamine release, complement activation in vitro, and occurrence of allergen-specific IgG. *J Allergy Clin Immunol* 82(2): 270-81.
- van Neerven, R. J., C. Ebner, H. Yssel, M. L. Kapsenberg and J. R. Lamb** (1996). T-cell responses to allergens: epitope-specificity and clinical relevance. *Immunol Today* 17(11): 526-32.
- van Oort, E., P. G. de Heer, W. A. van Leeuwen, N. I. Derksen, M. Muller, S. Huveneers, R. C. Aalberse and R. van Ree** (2002). Maturation of *Pichia pastoris*-derived recombinant pro-Der p 1 induced by deglycosylation and by the natural cysteine protease Der p 1 from house dust mite. *Eur J Biochem* 269(2): 671-9.
- Vernet, T., P. J. Berti, C. de Montigny, R. Musil, D. C. Tessier, R. Menard, M. C. Magny, A. C. Storer and D. Y. Thomas** (1995). Processing of the papain precursor. The ionization state of a conserved amino acid motif within the Pro region participates in the regulation of intramolecular processing. *J Biol Chem* 270(18): 10838-46.
- Vernet, T., H. E. Khouri, P. Laflamme, D. C. Tessier, R. Musil, B. J. Gour-Salin, A. C. Storer and D. Y. Thomas** (1991). Processing of the papain precursor. Purification of the zymogen and characterization of its mechanism of processing. *J Biol Chem* 266(32): 21451-7.
- Vernet, T., D. C. Tessier, C. Richardson, F. Laliberte, H. E. Khouri, A. W. Bell, A. C. Storer and D. Y. Thomas** (1990). Secretion of functional papain precursor from insect cells. Requirement for N-glycosylation of the pro-region. *J Biol Chem* 265(27): 16661-6.
- Visal, S., M. A. Taylor and D. Michaud** (1998). The proregion of papaya proteinase IV inhibits Colorado potato beetle digestive cysteine proteinases. *FEBS Lett* 434(3): 401-5.
- Volkel, H., U. Kurz, J. Linder, S. Klumpp, V. Gnau, G. Jung and J. E. Schultz** (1996). Cathepsin L is an intracellular and extracellular protease in *Paramecium tetraurelia*. Purification, cloning, sequencing and specific inhibition by its expressed propeptide. *Eur J Biochem* 238(1): 198-206.
- Voorhorst, R. and F. T. Spieksma** (1967). Correspondence: of mites and men. *J Allergy* 40(6): 357-8.
- Wan, H., H. L. Winton, C. Soeller, E. R. Tovey, D. C. Gruenert, P. J. Thompson, G. A. Stewart, G. W. Taylor, D. R. Garrod, M. B. Cannell and C. Robinson** (1999). Der p 1 facilitates transepithelial allergen delivery by disruption of tight junctions. *J Clin Invest* 104(1): 123-33.
- Wiederanders, B.** (2003). Structure-function relationships in class CA1 cysteine peptidase propeptides. *Acta Biochim Pol* 50(3): 691-713.
- Wiederanders, B., D. Bromme, H. Kirschke, K. von Figura, B. Schmidt and C. Peters** (1992). Phylogenetic conservation of cysteine proteinases. Cloning and expression of a cDNA coding for human cathepsin S. *J Biol Chem* 267(19): 13708-13.
- Yasueda, H., H. Mita, K. Akiyama, T. Shida, T. Ando, S. Sugiyama and H. Yamakawa** (1993). Allergens from *Dermatophagoides* mites with chymotryptic activity. *Clin Exp Allergy* 23(5): 384-90.
- Yasuhara, T., T. Takai, T. Yuuki, H. Okudaira and Y. Okumura** (2001). Cloning and expression of cDNA encoding the complete prepro-form of an isoform of Der f 1, the major group 1 allergen from house dust mite *Dermatophagoides farinae*. *Biosci Biotechnol Biochem* 65(3): 563-9.
- Zock, J. P., J. Heinrich, D. Jarvis, G. Verlato, D. Norback, E. Plana, J. Sunyer, S. Chinn, M. Olivieri, A. Soon, S. Villani, M. Ponzio, A. Dahlman-Hoglund, C. Svanes and C. Luczynska** (2006). Distribution and determinants of house dust mite allergens in Europe: the European Community Respiratory Health Survey II. *J Allergy Clin Immunol* 118(3): 682-90.