

Relationship between Propeptide pH Unfolding and Inhibitory Ability during ProDer p 1 Activation Mechanism

Andy Chevigné^{1,2}, Roya Barumandzadeh³, Sylvie Gros Lambert¹,
Benoît Cloes¹, Dominique Dehareng³, Patrice Filée¹,
Jean-Claude Marx¹, Jean-Marie Frère³, André Matagne³,
Alain Jacquet² and Moreno Galleni^{1*}

¹Macromolécules Biologiques,
Center for Protein Engineering,
Université de Liège,
Liège B-4000, Belgium

²Department of Applied
Genetics, Université Libre de
Bruxelles, Gosselies B-6041,
Belgium

³Laboratoire d'Enzymologie,
Center for Protein Engineering,
Université de Liège,
Liège B-4000, Belgium

Received 29 June 2007;
received in revised form
6 August 2007;
accepted 10 August 2007
Available online
21 August 2007

The major allergen Der p 1 of the house dust mite *Dermatophagoides pteronyssinus* is a papain-like cysteine protease (CA1) produced as an inactive precursor and associated with allergic diseases. The propeptide of Der p 1 exhibits a specific fold that makes it unique in the CA1 propeptide family. In this study, we investigated the activation steps involved in the maturation of the recombinant protease Der p 1 expressed in *Pichia pastoris* and the interaction of the full-length and truncated soluble propeptides with their parent enzyme in terms of activity inhibition and BIAcore interaction analysis. According to our results, the activation of protease Der p 1 is a multistep mechanism that is characterized by at least two intermediates. The propeptide strongly inhibits unglycosylated and glycosylated recombinant Der p 1 ($K_D = 7$ nM) at neutral pH. This inhibition is pH dependent. It decreases from pH 7 to pH 4 and can be related to conformational changes of the propeptide characterized by an increase of its flexibility and formation of a molten globule state. Our results indicate that activation of the zymogen at pH 4 is a compromise between activity preservation and propeptide unfolding.

© 2007 Elsevier Ltd. All rights reserved.

Edited by F. Schmid

Keywords: allergy; cysteine protease; inhibition mechanism; propeptide; pH unfolding

Introduction

House dust mite allergens are commonly associated with allergic diseases such as asthma, perennial rhinitis and atopic dermatitis. Der p 1 is a major allergen from *Dermatophagoides pteronyssinus* as more than 80% of the house dust mite-sensitized

population produce large amounts of Der p 1 specific IgE antibodies.^{1,2} It is a 25-kDa cysteine protease belonging to the papain-like protease family (CA1), which includes cathepsins K, L and B. Like most secreted cysteine proteases, Der p 1 is synthesized as a 34-kDa inactive precursor called ProDer p 1, consisting of a cysteine protease domain (222 residues) and an N-terminal propeptide (80 residues) that blocks the proteolytic activity. Each domain contains one potential N-glycosylation site in position 16p for the propeptide and in position 52 for Der p 1.^{3,4} The propeptide might also function as an intramolecular chaperone that ensures correct folding and targeting of the enzyme through the secretory pathway.

The proteolytic activity of Der p 1 has been reported to be an additional factor for the allergic

*Corresponding author. E-mail address:
mgalleni@ulg.ac.be.

Abbreviations used: AMC, 7-amino-4-methylcoumarin; ANS, 8-anilino-1-naphthalenesulfonic acid; Boc, N-t-butoxy-carbonyl; EDTA, ethylenediaminetetraacetic acid; PBS, phosphate buffered saline; ProDer p 1, protease Der p 1 zymogen; rDer p 1, recombinant Der p 1; SPR, surface plasmon resonance; WT, wild type.

pathogenesis.⁵ Several reports indicated that Der p 1 is able to increase bronchial permeability by disruption of the intercellular tight junctions and cleavage of the α 1-antitrypsin serine protease inhibitor.^{6,7} Furthermore, Der p 1 is capable of cleaving receptors, including the IgE low-affinity receptor (CD23) present at the surface of plasmocyte cells, the α -subunit of the interleukin-2 receptor (CD25) of T cells and the CD40 of dendritic cells.^{5,8-11} All these processes favor the development of the allergic response; thus, the development of specific Der p 1 inhibitors is of considerable interest.

Furthermore, the comprehension of the ProDer p 1 activation process is of considerable interest in the development of efficient recombinant mature allergen production systems thus essential for the screening of hypoallergenic vaccines.

Like most cysteine proteases, ProDer p 1 is activated under low pH conditions.¹²⁻¹⁴ Despite several attempts, *in vitro* demonstration of the activation of ProDer p 1 remains partly unsuccessful. For example, purified recombinant proforms of Der p 1 expressed in mammalian CHO and insect *Drosophila* cells could be converted into mature forms by incubation at pH 4 but do not exhibit enzymatic activity.^{15,16} More recently, however, several groups have succeeded in converting ProDer p 1 expressed in *Pichia pastoris* into mature enzymatically active forms with IgE reactivity.^{13,14} Finally, several experiments performed with ProDer p 1 mutants lacking the N-glycosylation sites or with endoglycosylase H-pretreated ProDer p 1 indicated that the glycosylation of recombinant zymogen expressed in *P. pastoris* can impair the maturation of the allergen.^{17,18}

Analysis of the X-ray structure of the zymogen form of the allergen reveals that the propeptide of Der p 1 adopts a new fold within the CA1 protease family¹⁹ (Fig. 1). This fold is characterized by the presence of an N-terminal globular domain formed by the association of three amphiphatic α -helices

maintained together by aromatic and hydrogen bond interactions. This domain, despite tertiary structure similarities, shows little or no conservation of the consensus residues characteristic of the cathepsin L-like propeptide subfamily.²⁰ Moreover, the presence of a fourth α -helix instead of a long terminal coil and a β -sheet at the C-terminal extremity make it unique among the previously described CA1 propeptides. Based on Karrer *et al.*'s classification, the Der p 1 propeptide could not be assigned to the cathepsins L and B subfamilies.²⁰ Meno *et al.* proposed that the Der p 1 propeptide belongs to a third subfamily characterized by a shorter propeptide with four α -helices and the absence of β -strands.¹⁹ Therefore, the study of the ProDer p 1 activation mechanism could represent an interesting way to highlight the importance of the presence or absence of a conserved propeptide motif and secondary structure elements during the activation of the papain-like precursor.

In the last decade, the potent inhibitor ability of several propeptides toward their corresponding enzyme of the papain-like family has been extensively studied.²¹⁻²⁴ Some cathepsin propeptides have been shown to display a strong inhibition capacity at neutral pH, with values of the dissociation constant K_D ranging from 0.12 nM for cathepsins L and B to 7.6 nM for cathepsins S and K. Interactions between the propeptides and the proteases are pH dependent: At neutral pH, slow binding interactions occur, whereas at acidic pH, weak or no interactions are observed.²¹ For the cathepsin L propeptide, lack of inhibition has been associated with partial unfolding of the propeptide and formation of a molten globule state under acidic conditions.²⁵ Moreover, under these conditions, it has been demonstrated that the cathepsin S propeptide is slowly degraded by mature cathepsin L.²¹

To date, the interaction between mature Der p 1 and its propeptide has not been investigated due to the difficulty of obtaining correctly matured

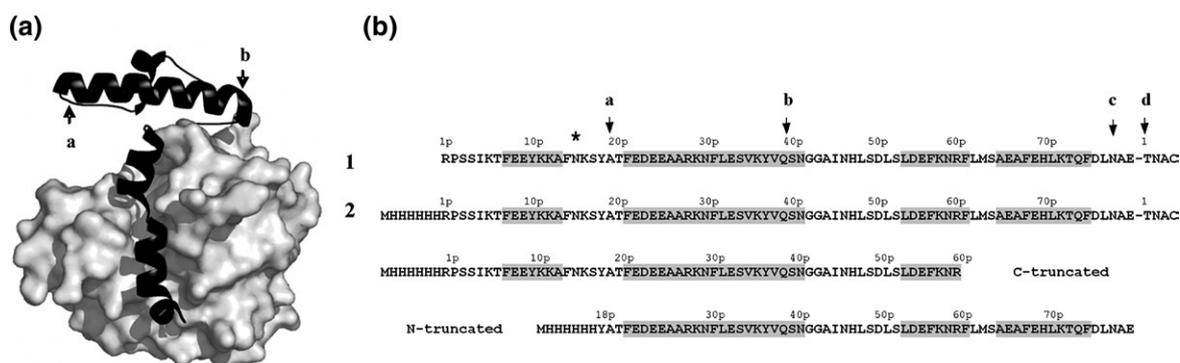


Fig. 1. Schematic representation of the different forms of the Der p 1 propeptide and the location of the cleavage sites identified during zymogen maturation. (a) Overall structure of the ProDer p 1 zymogen showing the unique propeptide folding (1XKG). The propeptide is shown in black, whereas the structure of the mature protease is shown in grey. (b1) Sequence of the full-length Der p 1 propeptide. The numbering of the propeptide is labeled with the letter *p*. Position 1 corresponds to the N-terminal amino acid of the mature Der p 1. Arrows *a*, *b*, *c* and *d* indicate the positions of the different cleavage sites of the propeptide during the maturation of the N52Q ProDer p 1 at pH 4. The α -helices are highlighted in grey. The position (16p) of the potential N-glycosylation site of the propeptide is indicated by an asterisk. (b2) Amino acid sequences of the His tag recombinant forms of the complete, C-truncated and N-truncated propeptides.

recombinant Der p 1 (rDer p 1). In order to characterize this interaction, we studied the activation steps involved in the zymogen maturation mechanism. Furthermore, we produced different recombinant forms of the Der p 1 propeptide (Fig. 1) and analyzed their interactions with mature glycosylated and unglycosylated rDer p 1 at different pH values in terms of activity inhibition and structural changes.

Results

Production and maturation of purified ProDer p 1 mutants

ProDer p 1 mutants encoding gene were constructed by overlapping PCR from the full-length ProDer p 1 synthetic cDNA with optimized mammalian codon usage.¹⁵ In these constructions, the potential N-glycosylation sites in the propeptide (N16p) and/or in the protease domain (N52) were eliminated by substitution of asparagine by glutamine (N16pQ, N52Q). Wild-type (WT) ProDer p 1 and the three mutants were successfully secreted by *P. pastoris*. Figure 2a indicates that the substitution of N16pQ in the ProDer p 1 yielded a mature glycosylated rDer p 1. Therefore, the hyperglycosylation site of the zymogen is the N52. This conclusion is supported by the fact that the rDer p 1 obtained by activation of ProDer p 1 N52Q or N16pQ/N52Q mutants is unglycosylated (Fig. 2a). However, surprisingly, for the N16pQ and N16pQ/N52Q mutants, partial activation occurred during the production, leading to the appearance of different

forms. The N-terminal sequencing of the purified mutants revealed the presence of at least three forms, previously described by Takai *et al.* and Jacquet *et al.*^{13,14} The first one displayed the ATFE sequence, which corresponds to the cleavage of the peptide bond between Tyr19p and Ala20p. This site is only three residues away from the propeptide N-glycosylation site and gives rise to the loss of the first α -helix of the propeptide (Fig. 1, site a). The two other forms showed the AETN and TNA sequences, which correspond to the mature allergen form with and without a two-residue extension, respectively. The ATFE intermediate could readily be converted into the mature allergen after incubation at pH 4. Such cleavages within the propeptide could not be observed for the WT proform produced under the same conditions, confirming previous observations that glycosylation of the propeptide might decrease the maturation ability of ProDer p 1.¹⁸

With the ProDer p 1 N52Q mutant, four bands (Fig. 2b, lanes 2 and 7) were observed at the end of the purification. In order to determine whether these bands correspond to partially activated ProDer p 1 or to a heterogeneous glycosylation at the N16p site, as proposed by Takai *et al.*, the proteins corresponding to the major bands (A and B) were analyzed by N-terminal sequencing.¹⁸ This revealed the presence of the complete proform (RPSS, band A) and the ATFE intermediate (band B). The N52Q mutant protein was submitted to the maturation assay under acidic conditions in both the presence and absence of 1 mM DTT and 1 mM ethylenediaminetetraacetic acid (EDTA) (Fig. 2b). Upon incubation in acidic buffer, an increase in the intensity of band C, with an apparent molecular mass of 27 kDa, was observed.

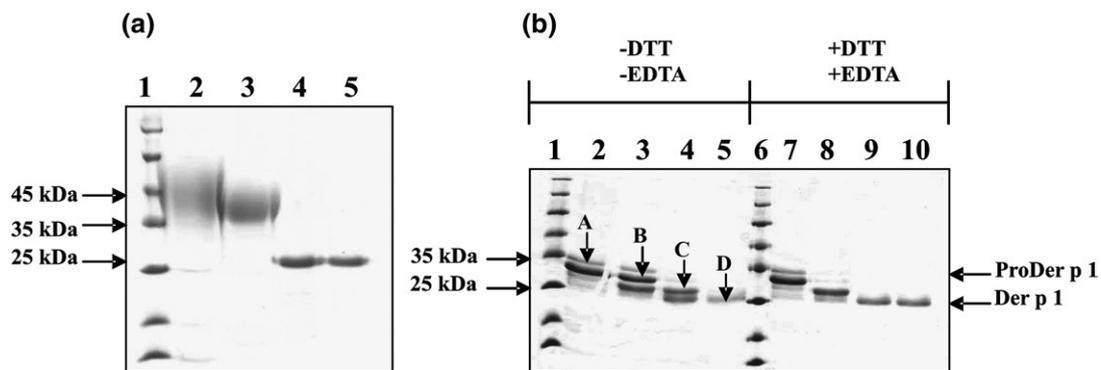


Fig. 2. Mutants glycosylation pattern and time-dependent processing of ProDer p 1 N52Q zymogen at pH 4. (a) Influence of glycosylation on the electrophoretic behavior of ProDer p 1 and Der p 1. Lane 1 represents the 10-kDa molecular mass marker (Fermentas). Lane 2 represents glycosylated WT recombinant ProDer p 1. Lane 3 represents glycosylated rDer p 1 yielded by the activation of N16pQ recombinant ProDer p 1. Lane 4 represents unglycosylated rDer p 1 yielded by glycosylated N52Q recombinant ProDer p 1. Lane 5 represents unglycosylated rDer p 1 (N52Q mutants) migrates faster and as sharper bands due to the absence of hyperglycosylation at the N52 site. (b) Identification of the different intermediates generated during ProDer p 1 activation. Purified N52Q mutant was incubated at 25 °C in polybuffer 50 mM (pH 4) with (right panel) or without (left panel) 1 mM DTT and 1 mM EDTA. Samples were analyzed by SDS-PAGE (Coomassie Blue staining). Lanes 1 and 6 represent 10-kDa molecular mass markers (Fermentas). Lanes 2 and 7 represent purified N52Q (i.e., time 0). Lanes 3 and 8 represent 2-h incubation. Lanes 4 and 9 represent 8-h incubation. Lanes 5 and 10 represent 24-h incubation. The N-terminal sequences of A, B, C and D are RPSS, ATFE, SNGG and AETNA/TNA, respectively.

N-terminal sequencing of band C revealed the occurrence of one additional intermediate species characterized by the SNGG sequence, which corresponds to the cleavage of the peptide bond between residues Gln40p and Ser41p of the propeptide (Fig. 1, site b). The evolution of the band intensities observed in the absence of DTT and EDTA (Fig. 2b) clearly indicates that this form appears after the ATFE intermediate. Finally, after 24 h of incubation, form C evolved in the mature form of Der p 1 (band D). These results seem to confirm that ProDer p 1 maturation is a multistep mechanism that leads to the formation of at least two intermediates. This process could be accelerated in the presence of DTT and EDTA (Fig. 2b), suggesting that the catalytic cysteine may play a major role in the activation of the zymogen.

Production and purification of the full-length and truncated propeptides

The ProDer p 1 His-tagged full-length (Arg1p-Glu80p), N-truncated (Tyr19p-Glu80p) and C-truncated (Arg1p-Arg60p) propeptides were produced in *Escherichia coli* as cytoplasmic soluble proteins. About 90 mg of purified proteins per liter of culture were obtained for the three propeptide forms. N-terminal sequencing of the three propeptides revealed the MHHH sequence, confirming the presence of the His-tag preceded by the N-terminal methionine. The molecular mass of the full-length form was determined by mass spectrometry, which gave a value of 10,227 Da, perfectly matching the value calculated from the amino acid sequence (10,227.3 Da).

Activity inhibition and pH dependence

Our kinetic experiments indicated that the full-length propeptide inhibits the proteolytic activities of unglycosylated (N52Q) and glycosylated rDer p

1, with dissociation constant values ($K_{D\text{Global}}$) of 7 ± 1 nM ($k_{\text{off}} = 0.008 \pm 0.001 \text{ s}^{-1}$) and 9 ± 4 nM ($k_{\text{off}} = 0.008 \pm 0.003 \text{ s}^{-1}$) at pH 7, respectively. Progress curves of product formation in the presence of increasing concentrations of unglycosylated propeptide (from 0 to 120 nM) are shown in Fig. 3a. These data highlight the slow binding interaction between the propeptide and rDer p 1 and allow calculation of the value of the pseudo first-order rate constant, k_{obs} , at various propeptide concentrations. K_D values were computed based on model 1 using Eq. (4) (Fig. 3b). With both glycosylated and unglycosylated rDer p 1, identical K_D values within the error limit were obtained, indicating that glycosylation of the protease domain does not impair the propeptide association. This observation can be explained on the basis of the position of the N52 glycosylation site, which, in the structure of the zymogen, is situated on the opposite side of the enzyme active site and therefore cannot interact with the propeptide during association. Interestingly, no significant inhibition could be detected with truncated propeptides (data not shown). At the highest concentration tested (5 μM), only a minor decrease ($\leq 10\%$) of the initial activity could be monitored with the N-truncated propeptide, whereas no change was detected with the C-truncated form. These observations show the importance of the C- and N-terminal domains of the propeptide for the inhibition process.

The activity of rDer p 1 and its interaction with the propeptide were shown to be pH dependent. The initial rates of product formation measured at different pH values are shown in Fig. 4a. rDer p 1 is active at pH values ranging from 3.5 to 9, with maximal activity at approximately pH 6.5. At pH 4, rDer p 1 retains 15% of its maximal activity, which can be sufficient to ensure zymogen activation. Our results are in good agreement with the pH activity profile established by Schulz *et al.*²⁶

A comparison of the steady-state rates of product formation measured at different pH values, in the

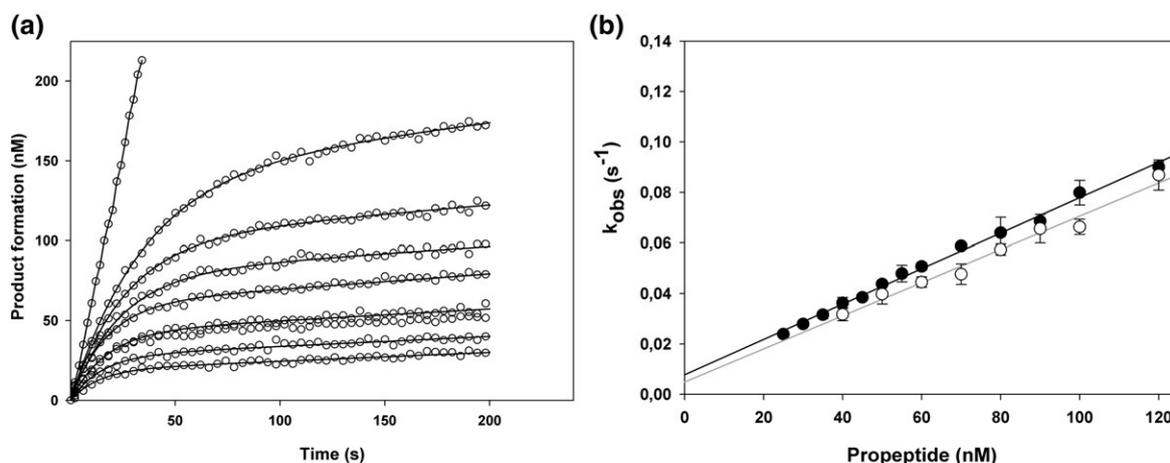


Fig. 3. Inhibition of rDer p 1 by its propeptide. Measurements were performed at 25 °C in 50 mM polybuffer (pH 7). (a) Time-dependent inactivation of rDer p 1 (5 nM) by propeptide (30–100 nM). The concentration of Boc-Gln-Ala-Arg-AMC was 160 μM . Each curve was obtained in triplicate. (b) Plot of k_{obs} versus propeptide concentration. The k_{obs} values were determined for the glycosylated (○) and unglycosylated (N52Q) (●) rDer p 1. Average values are given with the SD.

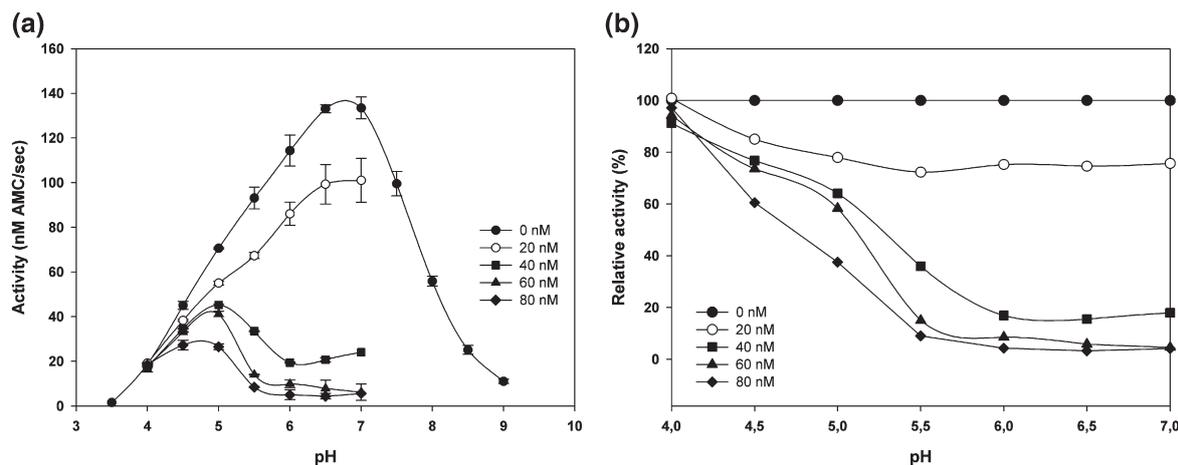


Fig. 4. pH dependence of the interaction between Der p 1 and its propeptide. Measurements were performed at 25 °C in 50 mM polybuffer between pH 3 and pH 9. (a) The activity of N52Q rDer p 1 (5 nM) at different pH levels and in the presence of propeptide (0 to 80 nM) was measured at the steady state of the hydrolysis rate of 40 μ M Boc-Gln-Ala-Arg-AMC. Each experiment was done in triplicate. Average values are given with the SD. (b) Dependence of the residual activity (%) of rDer p 1 in the presence of different propeptide concentrations.

presence of increasing propeptide concentrations, highlights the pH dependence of the inhibition process (Fig. 4a and b). At pH 4, the enzyme remains fully active even in the presence of high propeptide concentrations, whereas in the pH range 6–7, less than 20% of the enzyme activity is retained at a propeptide concentration of 40 nM or higher.

Propeptide/Der p 1 association and dissociation by surface plasmon resonance (SPR)

The affinity of the propeptide for rDer p 1 (N52Q) was first evaluated using a competition assay in solution, and a dissociation constant (K_D) value of 12 ± 2 nM could be determined (data not shown). This value indicates that the full-length propeptide displays a high affinity for Der p 1. In addition, a kinetic analysis allowed calculation of the rate constants for association (k_a) and dissociation (k_d). A summary of the calculated rate and equilibrium constants is shown in Table 1. A global analysis of the association and dissociation data with several models revealed that the data could not be fitted to a simple 1:1 binding model (model 1) but are better described either by a model that accounts for

conformational change (model 2), leading to the formation of a higher affinity complex (Fig. 5a), or by a heterogeneous ligand model (model 3) (Fig. 5b). The former model implies a 1:1 binding of rDer p 1 to the immobilized propeptide, followed by a conformational change of the complex, which can affect the propeptide, rDer p 1 or even both. The value of the global affinity constant obtained by fitting the data to model 2 ($K_D = 12.9 \pm 0.2$ nM) is not significantly different from the value obtained in solution. Alternatively, the heterogeneous ligand model yielded the K_D values of 7.2 and 76 nM. Such a model is often observed when the ligand has multiple attachment sites to the sensor chip. Two of the seven lysine residues, which participate to the covalent immobilization of the propeptide, are located in the C-terminal part of the propeptide. Thus, the lower affinity constant value (76 nM) could correspond to the propeptide molecule immobilized through these two residues, in a less accessible orientation for Der p 1. Even if these two models reflect two phenomena, the calculated constants for the two models are in the same order of magnitude and are in good agreement with those obtained in solution.

Table 1. Kinetic parameters

Constant	Two-state reaction	Heterogeneous ligand	Enzymatic inhibition	Simulation
k_{on} ($M^{-1}s^{-1} \times 10^4$)	–	–	110 ± 2.3	–
k_{off} ($s^{-1} \times 10^{-3}$)	–	–	7.8 ± 1.0	–
k_{a1} ($M^{-1}s^{-1} \times 10^4$)	10.1 ± 0.2	7.5 ± 0.02	–	7.5 ± 1.7
k_{d1} ($s^{-1} \times 10^{-3}$)	7.6 ± 0.6	0.5 ± 0.006	–	1.8 ± 0.05
k_{a2} ($s^{-1} \times 10^{-3}$) ($M^{-1}s^{-1} \times 10^5$ for the second column)	4.3 ± 0.1	1.6 ± 0.001	–	3.6 ± 0.05
k_{d2} ($s^{-1} \times 10^{-4}$)	9.0 ± 0.1	123 ± 0.87	–	9.0 ± 0.4
K_{Dinit} ($M \times 10^{-9}$)	75.2 ± 4.5	–	36.7 ± 1.8	24 ± 6
K_{Dfinal}	0.20 ± 0.01	–	0.23 ± 0.04	0.25 ± 0.08
$K_{Dglobal}$ ($M \times 10^{-9}$)	12.9 ± 0.2	–	7.0 ± 1.0	4.8 ± 1.1
K_{D1} ($M \times 10^{-9}$)	–	7.2 ± 0.1	–	–
K_{D2} ($M \times 10^{-9}$)	–	76.3 ± 1.2	–	–

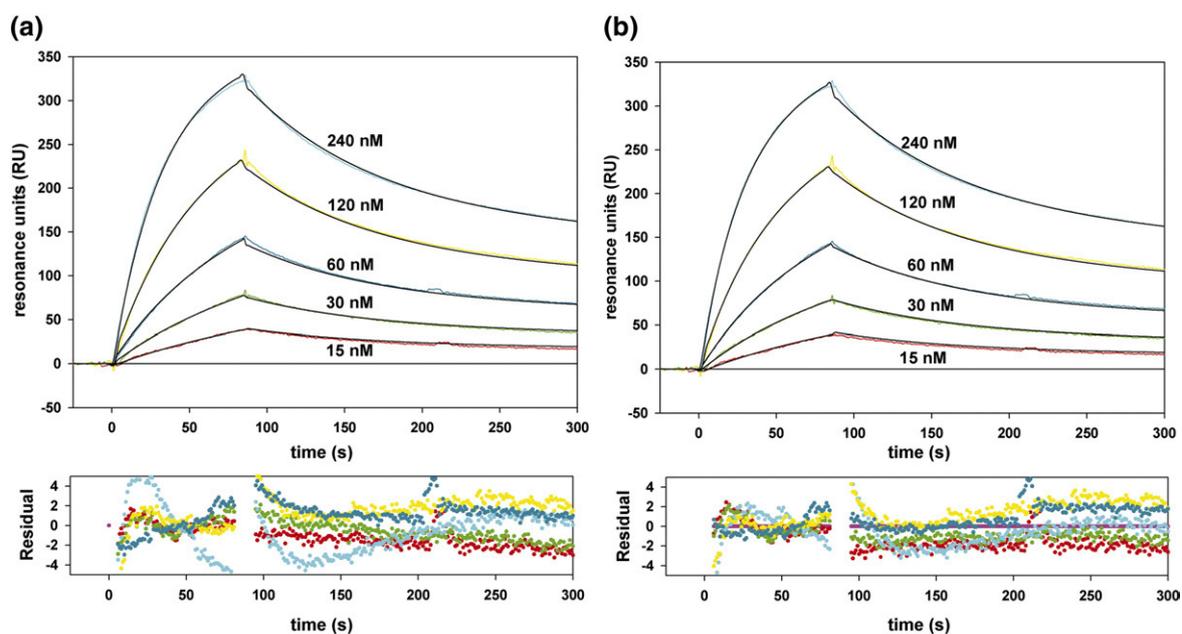


Fig. 5. Analysis by SPR spectroscopy of the binding of rDer p 1 to the full-length propeptide immobilized on a BIAcore CM5 chip. Different concentrations (0, 15, 30, 60, 120 and 240 nM) of rDer p 1 N52Q in Hepes buffered saline were injected at a flow rate of 70 $\mu\text{l}/\text{min}$. Kinetic rate constants (k_a and k_d) were fitted using: (a) the two-state reaction model including a conformational change and (b) the heterogeneous ligand model. The lower panel corresponds to the residual value of the fitting for each rDer p 1 concentration.

The validity of the two-state model was also analyzed by fitting the different rate constants determined by SPR to the different experimental product formation curves obtained at different propeptide concentrations. Primary fitting and optimization analysis of each time curve permitted determination of different sets of optimal rate constants for each initial propeptide concentration higher than 50 nM. The different rate constants

obtained from the primary analysis agreed fairly well with the values determined by SPR. Based on these data, the different sets of rate constants were averaged and recomputed for all product formation curves. The resulting optimized rate constants are listed in Table 1. Computed values permitted following the evolution of the concentration of the different enzyme species during the reaction time (Fig. 6a). Free enzyme concentration (E) decreases

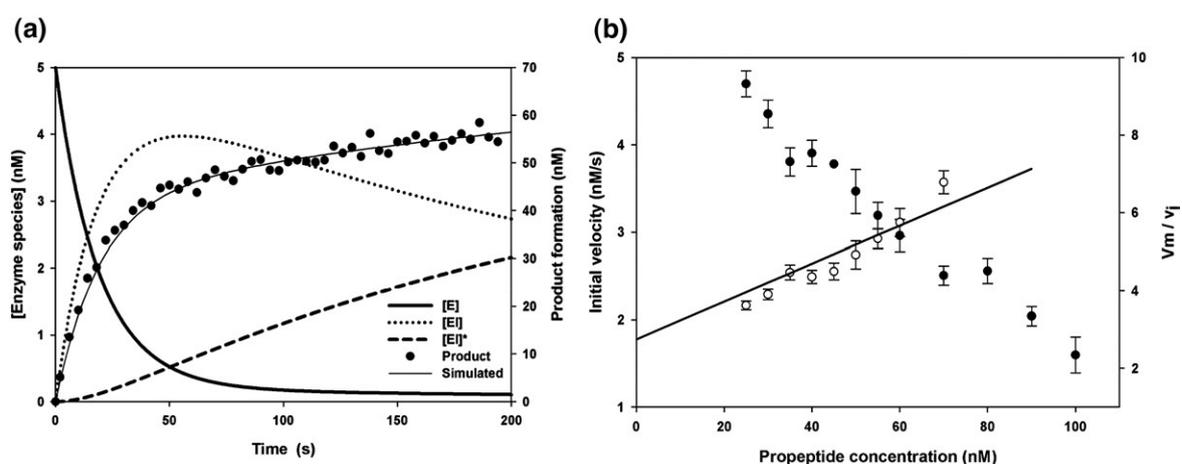


Fig. 6. Simulation of the evolution of the different complex species during the reaction time. (a) The left ordinate represents the evolution of the computed free enzyme [E] and complexes of species ([EI] and [EI*]), concentration *versus* time. The right ordinate represents experimental product formation curve measured in the presence of 70 nM propeptide and the corresponding simulated fitting curve. Data were computed from the experimental points recorded with an initial propeptide concentration of 70 nM using the optimized rate constants presented in Table 1. (b) The left ordinate represents the evolution of the initial rates *versus* initial propeptide concentrations. The right ordinate represents the evolution of the ratio of the maximal velocity over the calculated initial velocities as a function of the propeptide concentrations. Linear regression was performed using Grafit software.

rapidly as a well-shaped exponential curve, whereas the initial complex concentration (EI) increases in the first quarter of the reaction time and then decreases to reach about 55% of the $[E_0]$ at the end of the reaction time. The final complex (EI*) appears rapidly after EI, reflecting the rapid conversion of the initial complex in the more stable EI* complex. The final complex (EI*) concentration then increases to reach the remaining 45% of the initial enzyme concentration.

pH-Dependent conformational changes of the propeptide

The shift in the propeptide inhibition profile from pH 5.5 to pH 4 (Fig. 4b) could be explained by conformational changes of the propeptide. The near-UV circular dichroism (CD) spectrum of the propeptide at neutral pH shows two maxima of ellipticity at 278 and 286 nm, suggesting that at least one of the three tyrosine residues present in the N-terminal globular domain is in an asymmetric environment (Fig. 7a). Comparison of the near-UV CD spectra recorded from pH 7 to pH 2 reveals minor changes in the tertiary structure organization when pH shifts from 7 to 4, followed by a total loss of the near-UV CD signal between pH 4 and pH 2. On the other hand, far-UV CD measurements (Fig. 7b) indicate that the free propeptide is characterized by a high α -helical content, which is in good agreement with the analysis of the zymogen structure (Fig. 1). The far-UV CD spectra recorded at different pH values display no major change in the secondary structure at a pH of 4 or higher, whereas the spectra reveal a significant loss of the secondary structure content at pH 3 and pH 2. Changes in the near-UV CD spectra between pH 7 and pH 4 can be related to an increase in the intrinsic fluorescence of the tyrosine located in the N-terminal part of the propeptide. Furthermore, the pH dependence of the tyrosine fluorescence emission

was recorded at 305 nm (Fig. 8a). A significant increase in intensity is observed between pH 7 and pH 4, which coincides with the gentle decrease in CD signal observed at 278 nm. These changes in the optical properties of the propeptide can be explained by an increase in solvent accessibility of at least one tyrosine residue, with maximum exposure at pH 4 and a midpoint at pH 5.5. Taken together, these data suggest that acidification of the solution up to pH 4 leads to conformational changes in the globular N-terminal domain of the propeptide that are characterized by an increase in the solvent accessibility of at least one tyrosine residue and the conservation of a large part of the secondary and tertiary structures. At pH values lower than pH 4, the propeptide loses its tertiary and secondary structures. These results were confirmed by the study of the propeptide flexibility at different pH values (Fig. 8b). The minor increase of the value of the Stern–Volmer constant (K_{SV}) observed between pH 5 and pH 4 could be related to the conformational changes observed in CD and intrinsic fluorescence measurements, whereas the large increase of the K_{SV} value observed between pH 4 and pH 2 can be ascribed to the loss of tertiary and secondary structures observed in the near- and far-UV CD experiments, respectively (Figs. 8 and 9). All these data indicate that upon acidification to pH 4, the propeptide is transformed into a molten globule-like intermediate. Further evidence for the existence of this intermediate was obtained by monitoring the fluorescence of 8-anilino-1-naphthalenesulfonic acid (ANS) in the presence of the propeptide at different pH values. ANS fluorescence showed maximum intensity at pH values between 3 and 4, confirming the presence of a molten globule state under these pH conditions (Fig. 8c).

Finally, we could show that under acidic conditions (i.e., pH \sim 4), rDer p 1 rapidly degrades its propeptide into low-molecular-weight fragments (Fig. 9). This proteolysis could be associated with

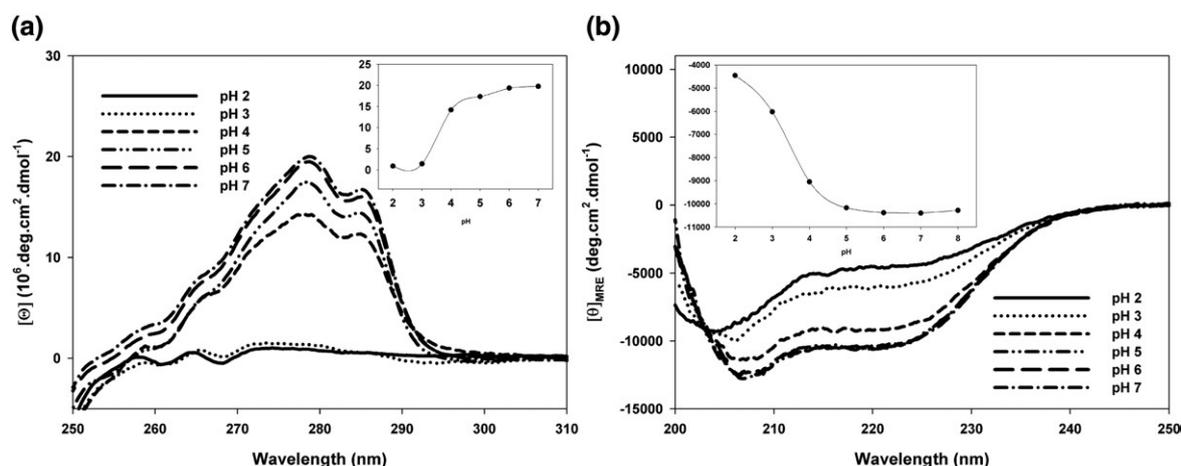


Fig. 7. Near- and far-UV CD spectra of full-length Der p 1 propeptide (100 μ M) at different pH values. Spectra were recorded at 25 $^{\circ}$ C in 10 mM polybuffer in the presence of 100 and 20 μ M of full-length propeptide for near- and far-UV CD spectra, respectively. The inset shows the pH-induced unfolding transition of the tertiary and secondary structure contents followed at 278 and 222 nm, respectively.

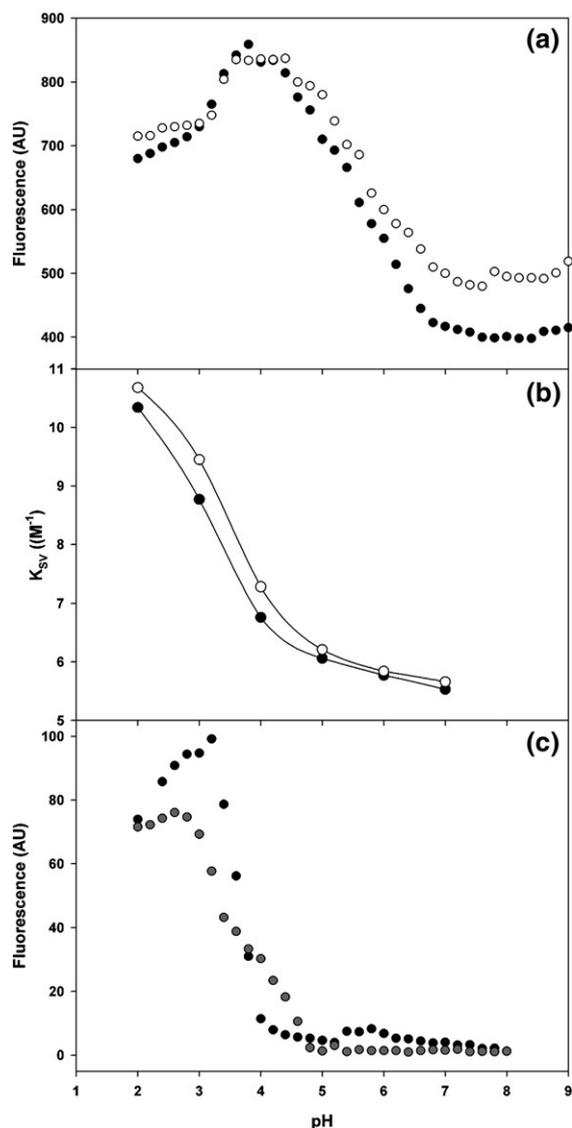


Fig. 8. Propeptide conformational changes followed by fluorescence spectroscopy. (a) Evolution of tyrosine residue intrinsic fluorescence recorded at 305 nm for full-length (●) and C-truncated (○) propeptides as a function of pH. Spectra were recorded at 25 °C in 50 mM polybuffer. The concentration of the propeptide was 5 μ M. (b) pH dependence of the acrylamide quenching constant K_{SV} values measured with the full-length (●) and C-truncated (○) propeptides. Fluorescence quenching was recorded at 305 nm in 50 mM polybuffer. (c) Evolution of ANS fluorescence recorded at 484 nm as a function of pH for full-length (●) and N-truncated (●) propeptides. Spectra were recorded at 25 °C in 50 mM polybuffer. Propeptide and ANS concentrations were 5 and 200 μ M, respectively.

the loss of the propeptide inhibition ability during incubation at pH 4 because no hydrolysis could be detected at pH values higher than 4.5. Similar experiments performed with rDer p 1 treated with E-64, a specific irreversible inhibitor of cysteine proteases, did not lead to propeptide hydrolysis even at pH 4, showing that the degradation of the

propeptide was due to the enzymatic activity of Der p 1.

Molecular modeling

Two average conformations of the propeptide were determined considering the time interval from 61.5 to 100 ps for the first one and from 601 to 1100 ps for the second. They were soaked in a 5-Å layer of water molecules, and the geometries of the whole systems were optimized. The two final geometries of the propeptide were compared with the starting one and superpositions were performed by taking into account all the heavy atoms of the propeptide (Fig. 10). It is clear that the major conformational changes from the propeptide geometry in the complex to its geometry as a free entity were occurring before 100 ps in the simulation because the two skeletons change much less in the final geometry than in the intermediate geometry. The geometry variation was the largest at the hinge between the α 2-helix and the α 3-helix, which mainly translates in a large change in two dihedral angles (p_i) defined by the C_α of four residues: residues 54, 53, 52 and 51 for p_1 and residues 55, 54, 53 and 52 for p_2 ; p_1 varies from -119 to -45 deg during the first 100 ps, whereas p_2 changes from -110 to -155 deg. The conformational changes result in the formation of a hydrophobic cluster at the (α 2- α 3) hinge as shown in Fig. 10. Furthermore, a significant change at the hinge between the α 3-helix and the α 4-helix also occurs after 100 ps.

Discussion

Previous studies highlighted difficulties to correctly mature recombinant ProDer p 1 expressed in

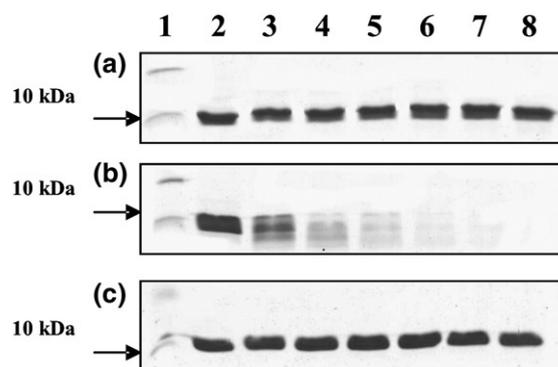


Fig. 9. pH-dependent proteolysis of the full-length propeptide by mature Der p 1. Time-dependent proteolysis of the propeptide at pH 7 (a) and pH 4 (b) in the absence (b) or presence (c) of E-64. A total of 70 μ g of propeptides was incubated at 37 °C with 3 μ g of mature N52Q rDer p 1 in 50 mM polybuffer containing 1 mM DTT and 1 mM EDTA at pH 7 and pH 4 and then analyzed by SDS-PAGE 18%. Lane 1 represents the 10-kDa molecular mass marker (Biorad, Hercules, USA). Lanes 2–8 correspond to 0.5, 2, 5, 10, 15, 30 and 40-min incubation periods of the propeptide in the presence of rDer p 1 N52Q.

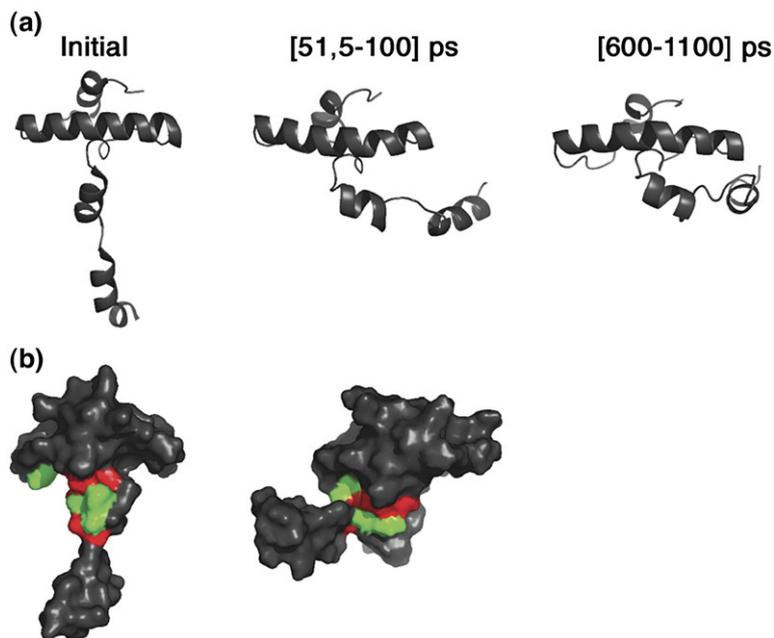


Fig. 10. Hydrophobic clustering of the hydrated free Der p 1 propeptide. (a) Initial geometry of the propeptide from 1XKG, intermediate geometry averaged on [51.5–100] ps and final geometry averaged on [601–1100] ps and then optimized. Only the skeleton of the C $_{\alpha}$ is drawn. (b) Detailed view of the (α 2– α 3) hinge showing the main hydrophobic residues that tend to form a cluster of the starting (left) and optimized [51.5–100 ps] (right) geometries. Red, Leu49p, Leu52p, Leu54p and Leu62p; green, Tyr38p, Phe57p and Phe61p.

different hosts and suggested that the glycosylation level of the zymogen ProDer p 1 could interfere with the activation process.^{17,18} In this study, we investigated the influence of the different potential N-glycosylation sites in the maturation process and the pH influence on the propeptide association.

Identification and localization of the cleavage sites in the activation of ProDer p 1

Elimination of the propeptide N-glycosylation motif present in the ProDer p 1 led to partial maturation of the corresponding zymogen during the fermentation process. This maturation was not observed for the WT ProDer p 1 expressed in the same conditions. Moreover, ProDer f 1, a homologous zymogen from *Dermatophagoides farinae*, which only possesses one N-glycosylation motif in its protease (N53) domain, is easily activated *in vitro*.²⁷ N-terminal sequencing of the bands obtained at the end of the fermentation for the N16pQ/N52Q ProDer p 1 mutant revealed the presence of a multistep activation mechanism initiated by the cleavage of the Tyr19p–Ala20p bond. The sequence of the first intermediate (ATFE) showed that the cleavage site is located in the solvent exposed coil that connects the first two α -helices of the propeptide (Fig. 1). On the other hand, the far-UV CD spectra of the full-length propeptide monitored at pH 4 showed that the propeptide retained a large part of its secondary structures at this pH value. This observation, together with the location of the cleavage site, indicated that the first cleavage occurred in a solvent-exposed part of the propeptide structure that does not possess a secondary structure at pH 4. Moreover, the N-truncated propeptide was unable to inhibit rDer p 1 even if present in a large concentration suggesting that cleavage of the first alpha helix could represent the limiting step of the activation mechanism. However, we cannot exclude

the possibility that this propeptide partially inhibits the allergen activity when still anchored to the enzyme. The fact that the formation of the ATFE intermediate, which is only three residues away from the propeptide N-glycosylation site, is the limiting step could explain the absence of partial activation observed for WT ProDer p 1. This observation is in good agreement with that of previous studies of the influence of glycosylation on the activation process. This hypothesis could also explain the different results reported by Jacquet *et al.* and van Oort *et al.*^{16,17} van Oort *et al.* did not succeed in obtaining maturation of WT recombinant ProDer p 1 with the correct N terminus,¹⁷ whereas Jacquet *et al.* reported that the WT ProDer p 1 lacking the first 19 residues was easily converted into mature allergen after acidic treatment.¹⁶ However, the fact that the ProDer p 1 N52Q mutant also presents the ATFE intermediate at the end of the production may be attributed to the heterogeneous glycosylation of the propeptide as previously proposed.¹⁸ Therefore, its presence could be attributed either to the partial activation of the minor glycosylated form or to the decrease of the steric hindrance corresponding to the lack of glycosylation of site 52 in the hypothesis of an intermolecular activation of the zymogen.

For the ProDer p 1 N52Q mutant, the maturation led to the appearance of a new band with a molecular mass intermediate between that of the ATFE form and that of mature allergen. This intermediate is characterized by the SNGG N-terminal sequence. This motif is located at the end of the second α -helix of the propeptide in a solvent-accessible area. Furthermore, this intermediate could correspond to the intermediate observed during the activation of the ATFE intermediate obtained with WT ProDer p 1.¹⁴ These two cleavages occurred in regions that correspond to sequences specifically recognized by Der p 1. Indeed, recent studies demonstrated that natural Der p 1 displays a predominant preference

for hydrophobic residues in position P2 and a slight preference for basic and polar amino acid residues in positions P1 and P3, respectively.^{26,28} Considering this substrate specificity, the two cleavage sites identified in the propeptide contained adequate residues at the P3-P2-P1 positions (i.e., Lys17p-Ser18p-Tyr19p and Tyr38p-Val39p-Gln40p), consistent with the Der p 1 specificity. N-terminal sequencing of the mature allergen forms indicated that cleavages also occurred at the end of the propeptide and generated the mature enzyme with a two-residue extension or with the correct N terminus. This phenomenon was also reported for other cysteine proteases and has been observed for ProDer p 1 maturation.^{12–14} These two cleavages occurred at sequences Asp76p-Leu77p-Asn78p and Asn78p-Ala79p-Glu80p and were in agreement with the Der p 1 substrate specificity. In other papain-like protease precursors, additional cleavages occurred in a region located just after the second β -strand.^{12,29} The lack of such cleavages in ProDer p 1 activation could be explained by the shorter length of the Der p 1 propeptide or by the presence of the fourth α -helix that could decrease the proteolysis accessibility of the corresponding area.

Propeptide inhibition

In this study, we investigated for the first time the interaction between soluble unglycosylated Der p 1 propeptide and its recombinant parent enzyme in terms of activity inhibition and SPR interaction analysis. The propeptide inhibits unglycosylated and glycosylated Der p 1 with K_D values of 7 and 9 nM, respectively. These dissociation constant values were confirmed using SPR technology. SPR permits the study of the association and the dissociation as two independent events that are complementary to the enzymatic approach. Fitting the SPR curves to different models allowed to show that two models could describe the interaction between rDer p 1 and its propeptide. From the comparison of the kinetic rate constants obtained from these two models and kinetic rate constants obtained by enzymatic analysis (Table 1), we could not exclude one of these two models. The dissociation constants determined by these two approaches were lower than the dissociation constants observed for the other papain-like propeptides but are similar to the results obtained for cathepsins K and L ($K_D \approx 2.6$ nM) for which the propeptide shows a high level of similarity with the Der p 1 propeptide.^{23,30} The presence of a two-step reaction mechanism had been proposed to explain the association of propeptides with cathepsin L.³⁰ This association could be initiated by weak contacts between the enzyme and the propeptide inducing conformational changes in one or both partners, leading to a more stable complex. Evidence of these conformational changes in the case of rDer p 1 could be demonstrated by the comparison of the initial and steady-state rates of the enzyme in the presence of increasing concentrations of the propeptide (Figs. 3a and 7b). Analysis of

the first part of the product formation curves showed an instantaneous decrease of the initial velocity when propeptide concentration increased, which indicated immediate inhibition and allowed the determination of the dissociation constant of 36.7 nM for the first equilibrium (75 nM for SPR). This transition complex was readily converted into a more stable complex with global K_D values of 7 nM for enzymatic analysis and 12.9 nM for SPR analysis.

The presence of a two-state interaction model was also analyzed and confirmed by fitting the experimental inhibition curves. Computer analysis permitted determination of the different rate constants describing the experimental curves by a two-equilibrium reaction model. Simulated values were in good agreement with those determined by SPR except for k_{a1} and k_{d1} , which are slightly different. This observation underlines the difficulty of comparing data recorded for the interaction between free partners or with one of them immobilized at the surface of a sensor chip. Analysis of the evolution of the two-complex concentration during the reaction time permitted to show the presence of initial and final complexes in the early reaction time, which could explain the difficulty of pointing out a two-state model in our kinetic inhibition analysis. Moreover, detection of a two-state model by inhibition analysis could be complicated by the presence of both inhibited complexes related by a rapid conversion rate ($K_{D\text{Final}}=0.2$).

As an illustration of the conformational changes between the initial (EI) and final (EI*) complexes, molecular modeling and dynamics showed that in water, the free propeptide undergoes fast conformational changes from its geometry in the complex that are mainly characterized by a rotation of the third and fourth α -helices in comparison with the N-terminal globular domain (Fig. 10). This conformational change could explain the observation of a two-equilibrium interaction model. Indeed, interactions could be initiated by loose contacts between the hydrated free propeptide forming the initial complex (EI) followed by a redocking of the fourth α -helix in the active site cleft driven by dehydration and the presence of a stereocomplementary charged area and hydrophobic patches leading to the formation of a more stable complex (EI*). In this state, the propeptide could adopt a conformation similar to that prevalent in the zymogen.

Free propeptide structure and pH unfolding

Previous studies of the interactions between proteases and propeptides did not intensively analyze the evolution of the structure of the free propeptide during acidification. Studies of the pH unfolding of the cathepsin L propeptide revealed that it exhibited a high content of secondary structures and a tertiary structure at neutral pH. On the other hand, at pH 4, it lost a large part of its tertiary structure but retained its secondary structure content.²⁵ In these conditions, the obtained propeptide intermediate was able to bind ANS. Our

data showed that, as for the cathepsin L propeptide, the Der p 1 propeptide and its C-truncated form were folded as compact globular domains at neutral pH. The ability of the free propeptide to adopt a correct fold independently of the protease domain demonstrates that this propeptide is a good candidate for acting as an intramolecular chaperone, as reported for other cysteine protease propeptides.^{24,31} No tertiary structure was detected by intrinsic fluorescence emission or by near-UV CD spectra for the N-truncated propeptide. This absence of a tertiary structure was due to the deletion of the first α -helix, which seems to play a key role in the folding of the propeptide.

Activation of the ProDer p 1 zymogen in the mature allergen must likely occur when interactions between the propeptide and the protease become weaker. For cysteine protease zymogens, this condition is reached by lowering the pH to 4. Proteolysis of the zymogen propeptide generates free propeptide or non-inhibitory propeptide fragments. Cleavages can result from intramolecular or intermolecular processes, depending on the zymogen.^{12,29} In the present study, we showed that the free propeptide of Der p 1 undergoes conformational changes under acidic conditions. At pH 4, this conformational transition led to a minor decrease of the tertiary structure but to a large increase in the solvent accessibility and flexibility of the residues located in the N-terminal globular domain. In this state, the propeptide retains a high amount of secondary structures and binds ANS. At lower pH, we have shown that the propeptide loses its residual tertiary and secondary structures and therefore its ability to bind ANS. No ANS binding was detected for the C-truncated propeptide, whereas increase in the ANS fluorescence was detected with the N-truncated form, suggesting that the C-terminal part of the propeptide plays a major role in the formation of the molten globule state. This may be due to the presence of the long fourth amphiphatic α -helix. The conformational changes recorded during acidification can be correlated to the shift of the inhibitory properties of the propeptide at acidic pH. Our results clearly show a relationship between the conformational changes recorded for the propeptide when pH decreases from pH 7 to pH 4 with a decrease of the interaction between the propeptide and the protease monitored by enzymatic inhibition

analysis. From these results, we can conclude that zymogen activation occurs at pH values inducing changes in the propeptide structure and permitting sufficient activity of the mature protease to enhance the activation process.

In conclusion, according to our results, the maturation process of ProDer p 1 under acidic conditions is a multistep process. In these conditions, the solvent accessibility of the propeptide residues increases and exposes the sequence specifically recognized by Der p 1. A first cleavage between Tyr19p and Ala20p of the propeptide liberates the first α -helix and thus further decreases the interaction between the two partners. This decrease results in rendering other parts of the propeptide, such as the end of the second α -helix and the C-terminal extremity of the propeptide, more sensitive to proteolysis.

Materials and Methods

Construction of the unglycosylated ProDer p 1 expression vectors

The N52Q, N16pQ and N16pQ/N52Q unglycosylated ProDer p 1 mutants (Der p 1 numbering) were constructed in two steps by overlapping PCR from pNIV4878.¹⁵ The sequences of the primers used in this study are detailed in Table 2. The amplified fragments were cloned into the pCR2.1 TOPO cloning vector (Invitrogen, Groenigen, The Netherlands), and the presence of both mutations of interest was verified by DNA sequencing. The fragments containing the ProDer p 1 coding sequence, isolated by digestion of the recombinant pCR2.1 TOPO with *Stu*-*Avr*II, were subsequently cloned into the pPIC9K expression vector (Invitrogen) previously restricted by *Sna*B1-*Avr*II. The resulting expression vectors contain the ProDer p 1 cassettes downstream of the *Saccharomyces cerevisiae* α -factor.

Expression of the ProDer p 1 proforms in *P. pastoris*

The recombinant expression vectors were introduced into *P. pastoris* SMD1168 (Invitrogen) using the electroporation transformation method. Transformants were first selected for histidinol dehydrogenase (His⁺) and then for increasing geneticin (G418) resistance (0.25 to 3 mg/ml). Selected clones were grown at 30 °C in 200 ml of yeast glycerol buffered media to an A_{600} value between 2 and 6.

Table 2. Nucleotidic sequences of the primers

Name	Sequence (5'–3')
ProNdeIUP	ATCCATATGCATCACCACCACCACCACCGCCGAGCTCCATTAAGACCTTCG
ProNdeI-18UP	ATCCATATGCATCACCACCACCACCACACTATGCCACCTTCGAGGACGAGGAGG
ProEcoRIDO	TCGCGAATTCCTACTCCCCTTGAGATCAAACCTGGGTC
ProEcoRI-20DO	TCGCGAATTCCTACTCCGTTCTTGAACCTCGTCTAAAGACAGGTC
N52QUP	CCTCGCGTATCGGCAACAGAGCCTGGACC
N52QDO	GGTCCAGGCTCTGTTGCCGATACGCGAGG
N16pQUP	ATACAAGAAAGCCTTCCAGAAGAGCTATG
N16pQDO	AGGTGGCATAGCTTCTTCTGGAAGGCTTCTTGTAT

Introduced restriction sites are in boldface.

This preculture was then transferred into 3.5 l of fermentation minimal medium and cultured for 24 h at 30 °C (pH 6). Production of the recombinant ProDer p 1 was then induced by the addition of methanol over 4 days (final concentration=0.5%). The methanol feed rate was regulated by following the dissolved oxygen level (30%). The supernatant was finally obtained by centrifugation of the culture at 13,000g for 20 min.

Purification and maturation of the ProDer p 1 proforms

The yeast culture medium containing the recombinant ProDer p 1 was dialyzed during 24 h against 20 mM Tris-HCl buffer (pH 8.5, starting buffer). The medium was then loaded onto a Q-HP Sepharose column (60 ml; Amersham Biosciences, GE Healthcare, Uppsala, Sweden) equilibrated with the starting buffer. The column was washed with the same buffer and the bound proteins were eluted with a linear NaCl gradient (0–0.3 M). The fractions containing ProDer p 1 were pooled and concentrated by ultrafiltration (cutoff=10 kDa). A homogeneous solution of ProDer p 1 could be obtained by gel filtration on a Sephacryl-100HR column (120 ml; Amersham Biosciences) equilibrated in 10 mM Na₂HPO₄/KH₂PO₄ buffer (pH 7.4) in the presence of 150 mM NaCl (phosphate buffered saline [PBS]). The concentration of ProDer p 1 was estimated by bicinchoninic acid assay (Pierce, Rockford, IL, USA).

ProDer p 1 maturation assay

The maturation was performed as previously described,¹⁴ with the following modifications: the purified ProDer p 1 was dialyzed against 100 mM acetate buffer at pH 4 during 72 h at 4 °C.^{13,14} For the N52Q proform, maturation was performed at 37 °C in polybuffer (mix of 50 mM Tris, phosphate, citrate, acetate and KCl, adjusted to chosen pH with HCl, pH 4), with or without the addition of 1 mM DTT and 1 mM EDTA. The maturation was followed by SDS-PAGE analysis, confirmed by N-terminal sequencing and enzymatic activity measurements.

Construction of the full-length and truncated propeptide expression vectors

The DNAs encoding to the full-length (Arg1p-Glu80p), N-terminal (Tyr19p-Glu80p) and C-terminal (Arg1p-Arg60p) truncated forms of the Der p 1 propeptide together with an N-terminal sequence encoding for six histidine residues (His-tag) were constructed by PCR using pNIV4878 as the template.¹⁴ The primers used to amplify the propeptide sequences were obtained from Eurogentec (Seraing, Belgium) (Table 2). The PCR products were cloned into pGEM-T Easy (Promega, Madison, WI, USA). The DNA fragments encoding the propeptide were sequenced to verify the correct amplification and the presence of the His tag. The propeptide coding cassettes were digested with NdeI and EcoRI restriction enzymes and cloned into pET-22b(+) (Novagen, Oxon, UK), previously restricted with the same enzymes.

Expression and purification of the propeptides

Recombinant expression vectors were transformed into *E. coli* BL21 (DE3) pLys cells (Novagen). Transformants

were selected on Luria-Bertani agar plates containing ampicillin (100 µg/ml) and chloramphenicol (30 µg/ml). The expression was initiated by a (1/50) dilution of an overnight subculture in 2 l of Luria-Bertani medium containing ampicillin (100 µg/ml) and chloramphenicol (30 µg/ml) until an *A*₆₀₀ value between 0.5 and 1 was reached. The propeptide expression was induced by the addition of IPTG (Immunosource, Halle-Zoersel, Belgium) to a final concentration of 1 mM in the medium. After 6 h of incubation at 28 °C, the cells were collected by centrifugation. The pellet was resuspended in PBS (pH 7.4) and then lysed with the help of a disintegrator (Constant Systems, Daventry, UK). The lysate was centrifuged for 20 min at 12,000g, and the pellet was discarded. The supernatant was applied onto a Ni-PDC column (10 ml; Affiland, Liège, Belgium) equilibrated in PBS. The column was washed with 10 volumes of PBS, and 5 volumes of the PBS 2 M NaCl. The propeptide was eluted with a linear imidazole (20–250 mM) gradient. The fractions enriched in the propeptide were pooled and dialyzed against 20 mM Tris buffer (pH 8). Further purification was achieved by anion-exchange chromatography on a MonoQ column (Amersham Biosciences) equilibrated with Tris buffer. The elution was performed with a linear NaCl (0–600 mM) gradient. Fractions containing the pure propeptide were pooled and dialyzed overnight against PBS. The homogeneity of the preparation was assessed by SDS-PAGE, N-terminal sequencing and mass spectrometry. The concentration of the purified propeptide was determined from the *A*₂₈₀ value, using the calculated extinction coefficient value of 3840 M⁻¹ cm⁻¹, and confirmed by a bicinchoninic acid assay (Pierce), using bovine serum albumin as standard. Purified soluble propeptides and processed ProDer p 1 were sequenced in an Applied Biosystems Procise 492 sequencer (Applied Biosystems, Foster City, USA), based on the Edman degradation.

Inhibition of the rDer p 1 by the propeptide

The proteolytic activity of rDer p 1 was estimated in the presence of various propeptide concentrations by following the increase in fluorescence emission resulting from the liberation of the 7-amino-4-methylcoumarin (AMC) group upon hydrolysis of the substrate Boc-Gln-Ala-Arg-AMC (Boc indicates *N*-t-butoxy-carbonyl; Bachem, Budendorf, Switzerland). The measured fluorescence intensity could be converted into product concentration, with the help of an AMC (0–0.6 µM) standard curve. Measurements were performed with an LS 50 B fluorimeter (Perkin Elmer, Wellesey, MA, USA) using excitation and emission wavelengths of 380 and 460 nm, respectively. Enzymatic reactions were initiated by the addition of active rDer p 1 (5 nM) to substrate solutions containing various concentrations of the propeptide. In all cases, the substrate concentration (40 or 160 µM) was below the *K*_m value (280 µM)²⁶ and the propeptide concentration (20–120 nM) was larger than the enzyme concentration. Under these conditions, it was possible to determine the value of the pseudo first-order rate constant (*k*_{obs}) characteristic of the complex formation and the individual kinetic rate constants for association (*k*_{on}) and dissociation (*k*_{off}). All experiments were carried out at 25 °C, between pH 9 and pH 3, in 50 mM polybuffer containing 1 mM DTT and 1 mM EDTA using a 1-ml stirred cell. The *k*_{obs} value was estimated from the nonlinear part of the product formation (P) time-course using the Grafit software (Erihtacus Software Ltd., Horley, UK). Thus, for each propeptide concentration, the apparent first-order rate constant (*k*_{obs})

was determined by fitting the decrease of the initial rate of the reaction (v_0) to its steady-state rate (v_s), according to Eq. (1), where k_{obs} depends on the propeptide concentration (I) according to Eq. (2) and considering Eq. (3):³²

$$p = v_s t + (v_0 - v_s) \frac{(1 - e^{-k_{\text{obs}} t})}{k_{\text{obs}}} \quad (1)$$

$$k_{\text{obs}} = k_{\text{off}} + k_{\text{on}} \left(\frac{K_m}{K_m + [S]} \right) [I] \quad (2)$$

$$K_{\text{DGlobal}} = \frac{k_{\text{off}}}{k_{\text{on}}} \quad (3)$$

According to model 1, the overall K_D (K_{DGlobal}) value was determined by linear regression of the first-order rate constant (k_{obs}) as a function of the propeptide concentration, according to Eq. (4):

Model 1:

$$E + I \xrightleftharpoons[k_{\text{off}}]{k_{\text{on}}} EI$$

$$k_{\text{obs}} = k_{\text{off}} \left(1 + \frac{[I]K_m}{K_{\text{DGlobal}}(K_m + [S])} \right) \quad (4)$$

The value of the initial dissociation constant (K_{DInit}) corresponding to the instantaneous inhibitory capacity of the propeptide was estimated by fitting the calculated initial rates (v_i) to Eq. (5) using the corresponding values of V_{max} (V_m), K_m and $[S]$:³²

$$\frac{V_m}{v_i} = 1 + \frac{K_m}{[S]} + \frac{[I]K_m}{K_{\text{DInit}}[S]} \quad (5)$$

The value of the final dissociation constant (K_{DFinal}) was determined using Eq. (6), considering the calculated values of the overall dissociation and initial dissociation:

$$K_{\text{DGlob}} = \frac{K_{\text{DInit}}K_{\text{DFinal}}}{K_{\text{DFinal}} + 1} \quad (6)$$

SPR (BIAcore) measurements

The interaction between the full-length propeptide and recombinant unglycosylated rDer p 1 was analyzed by SPR using a BIAcoreX instrument (BIAcore AB, Uppsala, Sweden). Binding under equilibrium conditions and the kinetics of binding were studied. Approximately 270 resonance units of full-length propeptide were immobilized on a carboxymethylated dextran-coated sensor chip CM5 (BIAcore AB) using the amine coupling chemistry (1-Ethyl-3-(dimethylaminopropyl)-carbodiimide/*N*-hydroxysulfosuccinimide) according to the manufacturer's instructions. The binding/regeneration cycles were performed at 25 °C in HEPES buffered saline containing EDTA and surfactant polysorbate 20 (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% v/v surfactant polysorbate 20). Surface regeneration was achieved by injection of 5 μ l of 6 M guanidinium chloride at a flow rate of 10 μ l/min. All measurements were repeated at least twice. In equilibrium binding experiments, unglycosylated rDer p 1 (128 nM) was incubated for 2 h (ensuring that the equilibrium was reached) with increasing concentrations of full-length propeptide (6 to 1500 nM), and then 100 μ l of the solution was injected on the CM5 surface coated with the propeptide at a flow rate of 30 μ l/min. The K_D value was determined using the model of "affinity in solution"

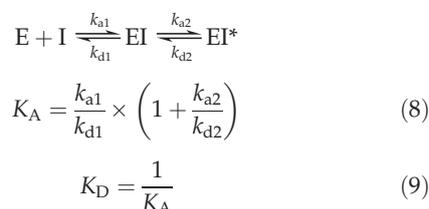
from the BIAevaluation 4.1 software, which is based on Eq. (7):

$$E_{\text{free}} = \frac{E - I - K_D}{2} + \sqrt{\frac{(I + E + K_D)^2}{4} - IE} \quad (7)$$

where E, E_{free} , I and K_D are the total concentration of rDer p 1, the concentration of unbound rDer p 1, the total concentration of the propeptide and the dissociation constant, respectively.

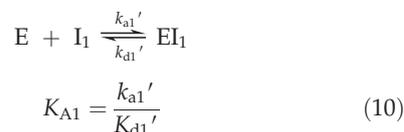
For kinetic experiments, different concentrations of unglycosylated rDer p 1 were injected at a flow rate of 70 μ l/min. At this flow rate, the mass transfer effects were found to be negligible. Kinetic rate constant values for association (k_a) and dissociation (k_d) (Table 1) were obtained on the basis of two models: a two-state reaction model (model 2) that assumes a conformational change or a heterogeneous ligand model (model 3) provided with the BIAevaluation software (version 4.1):

Model 2:

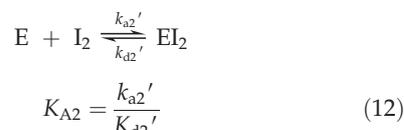


where I and E represent the propeptide and rDer p 1, respectively, k_{a1} and k_{d1} are the kinetic rate constants for association and dissociation of the initially formed complex EI, respectively, and k_{a2} and k_{d2} are the kinetic rate constants for interconversion between EI and EI*. K_A is the overall equilibrium association constant.

Model 3:



$$K_{D1} = \frac{1}{K_{A1}} \quad (11)$$



$$K_{D2} = \frac{1}{K_{A2}} \quad (13)$$

where E, I_1 and I_2 represent rDer p 1 and two propeptide conformations, respectively. The individual rate constants k_a and k_d correspond to the association and dissociation of the EI₁ and EI₂ complexes, respectively. K_{A1} and K_{A2} are the equilibrium association constants for each reaction.

Fluorescence measurements

All fluorescence experiments were performed with the help of a Perkin Elmer LS 50 B spectrofluorimeter using a stirred cell with 1 cm of path length. Propeptide intrinsic fluorescence and ANS-bound fluorescence emission spectra were recorded at 25 °C in 50 mM polybuffer (pH 2–9) using a scan speed of 280 nm per minute. The spectra were measured four times and averaged. Intrinsic fluorescence

measurements were performed using a protein concentration of 5 μM , with an excitation wavelength of 280 nm and emission recorded from 285 to 400 nm. The binding of ANS (Sigma, St. Louis, MO, USA) to the complete or truncated propeptides was measured using an excitation wavelength of 350 nm and recording emission spectra from 400 to 600 nm. These were corrected for the background fluorescence of free ANS. The concentration of ANS (200 μM) was determined using the value of the molar extinction coefficient of ANS at 350 nm ($4950 \text{ M}^{-1} \text{ cm}^{-1}$), and the $[\text{ANS}]/[\text{Propeptide}]$ ratio was about 40.

The dynamic quenching of fluorescence of tyrosine residues by acrylamide (ICN, Aurora, OH, USA) was monitored with excitation and emission wavelengths of 280 and 305 nm, respectively, and using a 50-s integration time. Propeptides were incubated with increasing concentrations of a small quencher molecule (acrylamide) to probe the accessibility of the tyrosine residues within the protein. The decrease of fluorescence intensity arising from diffusive collisions between the quencher and the fluorophore reflects the ability of the quencher to penetrate the structure and can be viewed as an index of protein flexibility.³³ The solvent accessibility of tyrosine residues was estimated by linear fitting of the (F_0/F) ratio, where F_0 and F are the fluorescence in the absence and that in the presence of acrylamide, respectively, according to the modified Stern–Volmer equation shown in Eq. (14), where K_{SV} represents the Stern–Volmer constant for dynamic quenching and $[Q]$ is the concentration of the quencher (i.e., acrylamide). The intrinsic protein fluorescence was corrected for the acrylamide inner filter effect (f), which was calculated using Eq. (15), with a molar extinction coefficient (ϵ) of $4.3 \text{ M}^{-1} \text{ cm}^{-1}$ for acrylamide at 280 nm.

$$\frac{F}{F_0} = 1 + K_{\text{SV}}[Q] \quad (14)$$

$$f = 10^{-\epsilon \frac{[Q]}{2}} \quad (15)$$

CD measurements

CD spectra were recorded using a Jasco J-810 CD spectropolarimeter (Jasco Inc., Easton, MD, USA) equipped with a Peltier temperature controller. Far- and near-UV CD spectra were recorded at 25 °C in 1 and 10 mm path length cells, respectively. Five scans were averaged and corrected by baseline spectra subtraction. All experiments were performed in 10 mM polybuffer (pH 2–7) using propeptide final concentrations of 25 and 100 μM for the far- and near-UV CD experiments, respectively. Spectra were acquired at a scan speed of 20 nm per minute, with a 1-nm bandwidth and a 4-s integration time.

Computer simulation

The enzymatic experimental data were fitted on the two-step model (model 2) using the equation set (Eqs. (16)–(20)) within the Matlab software environment (version 5.3, MathWorks Inc., Natick, USA). The routine uses the *lsqcurvefit* function, which solves nonlinear least squares problems. The differential equations of the model are solved by a low-order Runge–Kutta method (*ode23* function), which only needs to know the solution at the preceding point.

Because the values of the parameters slightly differ for the initial concentrations of the inhibitor, a unique set of values is chosen for the four parameters (the mean of the best fittings) and the product concentration (P) is simulated using the equation system (solved by a low-order Runge–Kutta method) in order to calculate the value of k_{obs} . The software gives the concentration of the different species versus time.

$$\frac{dE}{dt} = -k_1 \times E \times I + k_{-1} \times EI \quad (16)$$

$$\frac{dI}{dt} = -k_1 \times E \times I + k_{-1} \times EI \quad (17)$$

$$\frac{dEI}{dt} = -k_1 \times E \times I + k_{-2} \times EI^* - (k_{-1} + k_2) \times EI \quad (18)$$

$$\frac{dEI^*}{dt} = k_2 \times EI - k_{-2} \times EI^* \quad (19)$$

$$\frac{dP}{dt} = 0.36 \times k_{\text{cat}} \times E \quad (20)$$

where 0.36 represents the ratio of $[S]/K_m + [S]$.

Molecular modeling

The three-dimensional coordinates of the initial structure of the propeptide were taken from the crystal structure of the complex between Der p 1 and its propeptide¹⁹ and soaked in a 5-Å layer of water molecules. The geometry was first optimized with a convergence threshold of $0.02 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ and then submitted to two-dynamics simulations at 300K during 100 ps and 1.1 ns. The chosen force field was AMBER,^{34,35} the relative dielectric constant was set to 1 and the program used was Insight-Discover (Accelrys Software Inc.†) running on a Pentium 4 3.4 GHz under Unix Debian.

Acknowledgements

This work was supported by the Belgian National Fund for Scientific Research (FNRS), the Walloon Region (Direction Générale des Technologies, de la Recherche et de l’Energie) of Belgium (Vaccin 215133 and Allervac 616293) and the Fonds de la Recherche Fondamentale et Collective (2.4.524.03, 2.4.511.06 and 9.4519.98). A.C. is an aspirant of the FNRS (Brussels, Belgium). A.M. is a research associate of the FNRS and is supported in part by a grant from the Fonds de la Recherche Fondamentale et Collective (contract no. 2.4550.05).

We thank Marie-Eve Dumez for its helpful comments, Marc Vanhove for its helpful support on BIAcore analysis, Nicole Otthiers for performing the N-terminal sequencing and GlaxoSmithKline Belgium for providing access to ProDer p 1 DNA sequences.

References

1. Platts-Mills, T. A. & Chapman, M. D. (1987). Dust mites: immunology, allergic disease, and environmental control. *J. Allergy Clin. Immunol.* **80**, 755–775.

† <http://www.accelrys.com>

2. Chapman, M. D., Rowntree, S., Mitchell, E. B., Di Prisco de Fuenmajor, M. C. & Platts-Mills, T. A. (1983). Quantitative assessments of IgG and IgE antibodies to inhalant allergens in patients with atopic dermatitis. *J. Allergy Clin. Immunol.* **72**, 27–33.
3. Chua, K. Y., Stewart, G. A., Thomas, W. R., Simpson, R. J., Dilworth, R. J., Plozza, T. M. & Turner, K. J. (1988). Sequence analysis of cDNA coding for a major house dust mite allergen, Der p 1. Homology with cysteine proteases. *J. Exp. Med.* **167**, 175–182.
4. Chua, K. Y., Kehal, P. K. & Thomas, W. R. (1993). Sequence polymorphisms of cDNA clones encoding the mite allergen Der p 1. *Int. Arch. Allergy Immunol.* **101**, 364–368.
5. Kikuchi, Y., Takai, T., Kuhara, T., Ota, M., Kato, T., Hatanaka, H. *et al.* (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**, 1609–1617.
6. Wan, H., Winton, H. L., Soeller, C., Tovey, E. R., Gruenert, D. C., Thompson, P. J. *et al.* (1999). Der p 1 facilitates transepithelial allergen delivery by disruption of tight junctions. *J. Clin. Invest.* **104**, 123–133.
7. Brown, A., Farmer, K., MacDonald, L., Kalsheker, N., Pritchard, D., Haslett, C. *et al.* (2003). House dust mite Der p 1 downregulates defenses of the lung by inactivating elastase inhibitors. *Am. J. Respir. Cell Mol. Biol.* **29**, 381–389.
8. Shakib, F., Schulz, O. & Sewell, H. (1998). A mite subversive: cleavage of CD23 and CD25 by Der p 1 enhances allergenicity. *Immunol. Today*, **19**, 313–316.
9. Schulz, O., Sutton, B. J., Beavil, R. L., Shi, J., Sewell, H. F., Gould, H. J. *et al.* (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**, 584–588.
10. Schulz, O., Sewell, H. F. & Shakib, F. (1998). 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**, 271–275.
11. Furmonaviciene, R., Ghaemmaghani, A. M., Boyd, S. E., Jones, N. S., Bailey, K., Willis, A. C. *et al.* (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**, 231–242.
12. Menard, R., Carmona, E., Takebe, S., Dufour, E., Plouffe, C., Mason, P. & Mort, J. S. (1998). Autocatalytic processing of recombinant human pro-cathepsin L. Contribution of both intermolecular and unimolecular events in the processing of pro-cathepsin L in vitro. *J. Biol. Chem.* **273**, 4478–4484.
13. Takai, T., Mineki, R., Nakazawa, T., Takaoka, M., Yasueda, H., Murayama, K. *et al.* (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**, 265–272.
14. Jacquet, A., Magi, M., Petry, H. & Bollen, A. (2002). High-level expression of recombinant house dust mite allergen Der p 1 in *Pichia pastoris*. *Clin. Exp. Allergy*, **32**, 1048–1053.
15. Massaer, M., Mazzu, P., Haumont, M., Magi, M., Daminet, V., Bollen, A. & Jacquet, A. (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**, 32–43.
16. Jacquet, A., Haumont, M., Massaer, M., Daminet, V., Garcia, L., Mazzu, P. *et al.* (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**, 677–684.
17. van Oort, E., de Heer, P. G., van Leeuwen, W. A., Derksen, N. I., Muller, M., Huvneers, S. *et al.* (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**, 671–679.
18. Takai, T., Mizuuchi, E., Kikuchi, Y., Nagamune, T., Okumura, K. & Ogawa, H. (2006). Glycosylation of recombinant proforms of major house dust mite allergens Der p 1 and Der f 1 decelerates the speed of maturation. *Int. Arch. Allergy Immunol.* **139**, 181–187.
19. Meno, K., Thorsted, P. B., Ipsen, H., Kristensen, O., Larsen, J. N., Spangfort, M. D. *et al.* (2005). The crystal structure of recombinant proDer p 1, a major house dust mite proteolytic allergen. *J. Immunol.* **175**, 3835–3845.
20. Karrer, K. M., Peiffer, S. L. & DiTomas, M. E. (1993). Two distinct gene subfamilies within the family of cysteine protease genes. *Proc. Natl. Acad. Sci. USA*, **90**, 3063–3067.
21. Maubach, G., Schilling, K., Rommerskirch, W., Wenz, L., Schultz, J. E., Weber, E. & Wiederanders, B. (1997). The inhibition of cathepsin S by its propeptide—specificity and mechanism of action. *Eur. J. Biochem.* **250**, 745–750.
22. Groves, M. R., Taylor, M. A., Scott, M., Cummings, N. J., Pickersgill, R. W. & Jenkins, J. A. (1996). The prosequence of procaricain forms an alpha-helical domain that prevents access to the substrate-binding cleft. *Structure*, **4**, 1193–1203.
23. Billington, C. J., Mason, P., Magny, M. C. & Mort, J. S. (2000). The slow-binding inhibition of cathepsin K by its propeptide. *Biochem. Biophys. Res. Commun.* **276**, 924–929.
24. Wiederanders, B. (2003). Structure–function relationships in class CA1 cysteine peptidase propeptides. *Acta Biochim. Pol.* **50**, 691–713.
25. Jerala, R., Zerovnik, E., Kidric, J. & Turk, V. (1998). pH-Induced conformational transitions of the propeptide of human cathepsin L. A role for a molten globule state in zymogen activation. *J. Biol. Chem.* **273**, 11498–11504.
26. Schulz, O., Sewell, H. F. & Shakib, F. (1998). 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**, 222–224.
27. Takahashi, K., Takai, T., Yasuhara, T., Yokota, T. & Okumura, Y. (2001). Effects of site-directed mutagenesis in the cysteine residues and the N-glycosylation motif in recombinant Der f 1 on secretion and protease activity. *Int. Arch. Allergy Immunol.* **124**, 454–460.
28. Harris, J., Mason, D. E., Li, J., Burdick, K. W., Backes, B. J., Chen, T. *et al.* (2004). Activity profile of dust mite allergen extract using substrate libraries and functional proteomic microarrays. *Chem. Biol.* **11**, 1361–1372.
29. McQueney, M. S., Amegadzie, B. Y., D’Alessio, K., Hanning, C. R., McLaughlin, M. M., McNulty, D. *et al.* (1997). Autocatalytic activation of human cathepsin K. *J. Biol. Chem.* **272**, 13955–13960.
30. Guo, Y. L., Kurz, U., Schultz, J. E., Lim, C. C., Wiederanders, B. & Schilling, K. (2000). The alpha1/2 helical backbone of the prodomains defines the

- intrinsic inhibitory specificity in the cathepsin L-like cysteine protease subfamily. *FEBS Lett.* **469**, 203–207.
31. Muntener, K., Willimann, A., Zwicky, R., Svoboda, B., Mach, L. & Baici, A. (2005). Folding competence of N-terminally truncated forms of human procathepsin B. *J. Biol. Chem.* **280**, 11973–11980.
 32. Morrison, J. F. & Walsh, C. T. (1988). The behavior and significance of slow-binding enzyme inhibitors. *Adv. Enzymol. Relat. Areas Mol. Biol.* **61**, 201–301.
 33. Lakowicz, J. (1983). *Principles of fluorescence spectroscopy*, Plenum Press, New York.
 34. Weiner, P. K. & Kollman, P. A. (1981). AMBER: Assisted Model Building with Energy Refinement. A general program for modeling molecules and their interactions. *J. Comput. Chem.* **2**, 287–303.
 35. Weiner, S. J., Kollman, P. A., Nguyen, D. T. & Case, D. A. (1986). An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **7**, 230–252.