Molecular characterisation of major allergens from mite (Lep d 2) and cat (Fel d 1)

By Liselotte Kaiser

Stockholm 2004
All previously published papers were reproduced with permission from the publisher.

Published and printed by intellecta docusys.
© Liselotte Kaiser, 2004
ISBN 91-7349-859-9

Med en stoisk gest kastar sig Cato på sitt svärd.
Jag föredrar att lugnt gå på en båt.

...ge mig av till sjöss och se litet närmare på havssidan av världen. Det är ett sätt att driva bort min spleen och stimulera blodomloppet. När jag får ett bistert drag kring munnen, när det råder november i min själ, // Då anser jag det vara hög tid att gå till sjöss så fort jag kan.

Ur Moby Dick av Herman Melville
översatt av Hugo Hultenberg
Abstract

The prevalence of allergic disease has increased significantly during the last decades and today more than 30% of the Swedish population is affected. By using recombinant allergens, diagnostics and treatment can be improved and the mechanisms of allergic reactions may be further explored. The objective of this thesis was to characterise allergens from two of the most important sources of allergens world-wide, cats and mites, at the molecular level. (Study I) Traditionally, mites grown in culture for years have been the source of allergens for use in research, diagnostics and therapy. However, the actual source of mite sensitisation is wild mites. The aim of this study was to investigate the occurrence of polymorphism in the gene encoding the allergen Lep d 2 in the dust mite Lepidoglyphus destructor derived from different sources. We could identify two Lep d 2 genes in both cultured mites and wild mites isolated from a hay sample. In addition, two new variant forms of Lep d 2 with single amino acid exchanges were found at different frequencies in wild and cultured mites. Neither in ELISA inhibition studies nor in T-cell experiments could any major differences between the variants be detected. (Studies II, III) To increase the understanding of the mechanisms underlying cat allergy and to obtain knowledge for new treatment strategies, we have solved the crystal structure of the most potent cat allergen, Fel d 1. The structure of Fel d 1 was determined by X-ray crystallography at 1.8 Å resolution. The fold of Fel d 1 presents a striking resemblance to uteroglobin, a molecule with anti-inflammatory and immunomodulatory properties. This resemblance provides possible explanations for the allergenicity of Fel d 1. An internal pocket was found within Fel d 1, and residual electron density within the cavity indicates the presence of a ligand. The shape of the cavity as well as the properties of the residues lining it indicate an amphipathic ligand. Finally, the surface localisation of three previously defined Fel d 1 IgE epitopes is presented. (Study IV) Today, the only curative treatment for allergic disease is immunotherapy. Although this treatment often results in reduced allergic symptoms, a major disadvantage is the risk of side effects. To reduce these risks, allergens with decreased IgE-binding capacity, but retained T-cell reactivity (hypoallergens) can be used. In this study, we present a new approach of how to generate hypoallergens using structural information and knowledge of B- and T-cell epitopes. The structure of the model allergen Fel d 1 was systematically altered by duplication of T-cell epitopes. In addition, disruption of disulphide bonds was performed. A panel of seven Fel d 1 derivatives was generated, and three of the derivatives displayed a marked reduction in IgE-binding capacity. In addition, they induced a lower degree of basophil activation and similar or stronger T-cell proliferation than Fel d 1. In conclusion, the work presented in this thesis exemplifies how detailed molecular analyses of allergens may provide knowledge about allergen function and offer tools for improved diagnosis and treatment of allergic disease. Specifically, the Lep d 2 sequence diversity identified in cultured and wild L. destructor mites seemed not to have any significant impact on IgE-binding or T-cell response. The determination of the crystal structure of Fel d 1 has provided clues about the pathogenesis of cat allergy as well as tools for improved diagnostics and therapy. Moreover, the new strategy to generate hypoallergens may be applied to any allergen to improve immunotherapy.
List of publications

This thesis is based on the following articles, which will be referred to by their roman numbers:


# CONTENTS

1 List of abbreviations ............................................................................................................ 7

2 Introduction ...................................................................................................................... 8
   2.1 The immune system .................................................................................................. 8
   2.2 Allergy .................................................................................................................... 8
   2.3 Allergens ................................................................................................................ 10
      2.3.1 General aspects of allergens ........................................................................... 10
      2.3.2 Allergenicity ................................................................................................... 11
      2.3.3 Recombinant allergens ................................................................................. 11
      2.3.4 Allergen 3D structure .................................................................................... 12
   2.4 Mite allergy ............................................................................................................. 15
      2.4.1 Sensitisation to mites and mite allergens ....................................................... 15
   2.5 Cat allergy ............................................................................................................. 18
      2.5.1 Sensitisation to cat and cat allergens ............................................................... 18
   2.6 Diagnosis and treatment of allergic patients ......................................................... 20
      2.6.1 Current diagnostics and therapy ..................................................................... 20
      2.6.2 Future diagnostics and therapy ....................................................................... 21

3 Aims of the thesis ............................................................................................................. 23

4 Material and methods ..................................................................................................... 24
   4.1 Mites ....................................................................................................................... 24
   4.2 Serum samples ......................................................................................................... 24
   4.3 PCR amplification and cloning ............................................................................... 24
   4.4 Site directed mutagenesis ....................................................................................... 24
   4.5 Genomic DNA extraction, southern blot analysis ................................................... 25
   4.6 Protein expression and purification ......................................................................... 25
   4.7 ELISA ..................................................................................................................... 25
   4.8 PBMC stimulation .................................................................................................... 25
   4.9 Basophil activation ................................................................................................... 26
   4.10 Fel d 1 crystallisation .............................................................................................. 26
   4.11 Data collection and processing ............................................................................... 26
   4.12 Phasing, model building and refinement .................................................................. 26

5 Results and discussion ..................................................................................................... 27
   5.1 Lep d 2 polymorphisms in *Lepidoglyphus destructor* mites (I) ............................. 27
   5.2 Determination of the 3D structure of Fel d 1 (II, III) .............................................. 32
   5.3 Rational design of hypoallergens (IV) ..................................................................... 38

6 Future perspectives ......................................................................................................... 42

7 Populärvetenskaplig sammanfattning ............................................................................. 43

8 Acknowledgements ....................................................................................................... 45

9 References ....................................................................................................................... 46
## 1 LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D</td>
<td>three-dimensional</td>
</tr>
<tr>
<td>APC</td>
<td>antigen presenting cells</td>
</tr>
<tr>
<td>CD</td>
<td>cluster of differentiation</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CRD</td>
<td>component-resolved diagnostics</td>
</tr>
<tr>
<td>D</td>
<td>duplicated</td>
</tr>
<tr>
<td>DTE</td>
<td>duplication of T-cell epitope</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>IFN</td>
<td>interferon</td>
</tr>
<tr>
<td>Ig</td>
<td>immunoglobulin</td>
</tr>
<tr>
<td>IL</td>
<td>interleukin</td>
</tr>
<tr>
<td>IUUS</td>
<td>the International Union of Immunological Societies</td>
</tr>
<tr>
<td>mAb</td>
<td>monoclonal antibody</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>MPD</td>
<td>2-methyl-2,4-pentanediol</td>
</tr>
<tr>
<td>PBMC</td>
<td>peripheral blood mononuclear cells</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>protein data bank</td>
</tr>
<tr>
<td>r</td>
<td>recombinant</td>
</tr>
<tr>
<td>RAST</td>
<td>radio allergosorbent test</td>
</tr>
<tr>
<td>rmsd</td>
<td>root mean square deviation</td>
</tr>
<tr>
<td>SeMet</td>
<td>seleno-methionine</td>
</tr>
<tr>
<td>SI</td>
<td>stimulation index</td>
</tr>
<tr>
<td>SIT</td>
<td>allergen-specific immunotherapy</td>
</tr>
<tr>
<td>TH</td>
<td>T-helper</td>
</tr>
</tbody>
</table>
2 INTRODUCTION

2.1 THE IMMUNE SYSTEM
The immune system is the human body's defence against potentially harmful pathogens. It consists of two parts. The first line of defence is almost as old as the first multi-cellular organism and is termed the innate or unspecific immune system. Protection against infections is sustained by a number of different cell types such as phagocytes and the natural killer cells, but also by proteins such as anti-bacterial peptides and complement factors. The cells of the innate immune system react immediately when common patterns of pathogens bind to pattern recognition receptors on their surface. The innate immune system generates no immunological memory. The second line of defence is the adaptive immune system, which is dependent on the ability of immunoglobulin (Ig)-like genes to undergo rearrangement, giving rise to gene diversity, a phenomenon found in vertebrates. This results in the lymphocyte T- and B-cell receptors (capable of recognising any molecule). If T- or B-cells carrying these receptors recognise peptides or proteins originating from pathogens they are activated, but only if the T-cell is simultaneously stimulated by co-stimulatory molecules on professional antigen presenting cells (APC), or B-cells by T-helper (H) cells. After activation and several days of clonal expansion naive lymphocytes differentiate into armed effector T-cells and antibody-producing B-cells, leading to the establishment of an immunological memory. When a harmless non-pathogenic antigen encounters the immune system two types of responses are possible. In most cases no effector mechanisms are activated since the host has developed tolerance against that molecule. However, in some individuals an immune response is induced, which may result in allergy. (Janeway et al. 2001)

2.2 ALLERGY
Von Pirquet coined the term allergy already in 1906 to describe a disease that affects the body after contact with antigens/allergens. Allergic disease has been rare until the last decades, since when the prevalence has increased dramatically in western societies (Lundbäck 1998, Beasley et al. 2000). Several factors have been proposed as important for the development of allergic disease, among them a family history of allergic disease (Wright 2004). In addition, "the hygiene hypothesis" coined in 1989 proposes that changes in family size and cleaner homes during the 20th century explain the increase in allergic disease (Strachan 1989). Although there is support for the hygiene hypothesis, there is currently no definitive proof that a reduced microbial burden is the cause of the increasing prevalence of allergy.

Atopy is a personal or familial tendency to produce IgE antibodies in response to low doses of allergens, usually proteins, and to develop allergic symptoms such as asthma, rhinoconjunctivitis or eczema/dermatitis (Johansson et al. 2001). The main characteristic of IgE-mediated allergy, previously referred to as Type I allergy (Coombs and Gell 1975), is the occurrence of allergen-specific IgE antibodies (Ishizaka et al. 1966, Johansson and Bennich 1967). IgE is found in low concentrations in serum.
and bound to the high affinity IgE receptors on mast cells, basophils, activated eosinophils and to a varying extent on APC.

Sensitisation is the initial event in the development of IgE-mediated allergic disease. During the sensitisation phase, allergen is taken up by APC and processed into peptides. The peptides are thereafter presented by the APC on the major histocompatibility complex (MHC) class II molecules to allergen-specific T-cells, leading to activation of the T-cells. The T-cells may subsequently activate allergen-specific B-cells (Kay 2001). In atopic individuals the allergen-specific T-cells are polarised to a Th2 type, which secrete interleukin (IL)-4 and IL-13, both of which promote antibody class switch to an IgE isotype in B-cells, and IL-5, which amplifies eosinophil production and activation (Imada et al. 1995, Till et al. 1997). Subsequently, IgE antibodies secreted by B-cells will bind to high affinity IgE receptors on the surface of mast cells.

When a sensitised individual is re-exposed to the same allergen, the allergen will bind to the IgE antibodies that are bound to the mast cells. Once IgE antibodies are cross-linked by the allergen, the mast cells will release mediators such as histamine and prostaglandins. These mediators are responsible for the immediate reactions, e.g. itching, swelling, bronchoconstriction and mucus secretion. The late phase reaction starts 6-9 hours after activation of the mast cells by synthesis and release of e.g. leukotrienes, chemokines and cytokines, which results in infiltration of inflammatory cells such as eosinophils and T-cells. The late phase reaction is followed by smooth muscle contraction, edema, airway hyperreactivity and asthma (Kay 2001). In figure 1 the allergic reaction is summarised. Although the pathological mechanisms of allergy have been thoroughly studied, the mechanisms underlying the unresponsiveness to allergens seen in healthy individuals are not clear. However, the role of regulatory cells that suppress Th2 responses have been discussed and a number of regulatory T-cell populations have been identified including Th3-cells, T-regulatory-1 cells and cluster of differentiation (CD)4+CD25+ T-cells (Umetsu et al. 2003). Interestingly, it has recently been shown that CD4+CD25+ T-cells from non-atopic individuals suppress proliferation of allergen-specific CD4+CD25+ T-cells, in contrast to CD4+CD25+ T-cells from atopic individuals (Cavani et al. 2003, Ling et al. 2004).

Atopy and asthma are multi-gene diseases and the genetic background is complex. Nevertheless, a number of candidate genes for allergy and asthma have been identified (Cookson 2002). The loci most consistently identified are on chromosome 5, 6, 12 and 13. These chromosomes contain several genes encoding proteins that are involved in the development of allergic disease: IL-4, IL-13, IL-5, interferon (IFN)-γ, granulocyte-macrophage colony-stimulating factor, CD14, the MHC gene cluster, the gene for the high affinity IgE receptor, genes linked to the regulation of IgE and IgA levels and the ADAM33 gene, which has been associated with asthma (Cookson 2002).
2.3 ALLERGENS

2.3.1 General aspects of allergens

Allergens are common proteins or glycoproteins that are innocuous to most individuals, except for those who have a predisposition for allergic sensitisation. Most allergens have a molecular weight of 10-100 kDa and are highly soluble and stable proteins (Valenta et al. 2004). Aqueous allergen extracts have been studied extensively for almost half a century and individual allergens have been purified and characterised. To manage all documented allergens, the International Union of Immunological Societies (IUIS) Allergen Nomenclature Sub-Committee maintain a database at http://www.allergen.org, which today contains information on about 470 allergens, though it is far from complete. A detailed description of different allergen databases is given in a recent review (Brusic et al. 2003).

The sources from which allergens come contain a large number of components, both allergenic and non-allergic. The allergenic components, the allergens, are designated as follows: the first three letters of the genus of the source, the first letter of the species followed by an Arabic number (WHO/IUS 1995). For example, Lep d 2 refers to an allergen identified from the dust mite Lepidoglyphus destructor. In nature, there can be several isoforms of an allergen due to the existence of different alleles of the allergen gene, encoding allergens with slightly different amino acid sequences. Isoforms are therefore very similar but may exhibit minor differences in three dimensional (3D) structure, size and biological function (Nedergaard Larsen 1995). These differences can also affect the B- and T-cell epitopes of the allergens and some isoforms are found to be more immunogenic than others (Ferreira et al. 1996). By using two-dimensional gel electrophoresis Le Mao and co-workers have identified up to ten isoforms of a major allergen from Dermatophagoides farinae (D. farinae), Der f 1, which were recognised by IgE (Le Mao et al. 1998). Isoallergens with closely similar sequences, > 97%
sequence identity, are designated allergen variants. A four-digit number designates the isoallergens and the corresponding variants. The first two figures refer to the isoallergen and the following two to the variant. The digits are chosen in the order of identification (WHO/IUS 1995). For example, Lep d 2.0201 refers to the first allergen variant identified belonging to the second isoform of the Lep d 2 allergen.

### 2.3.2 Allergenicity

To recognise the properties that make a protein an allergen is important, not only to understand the development of allergy but also for risk assessment of industrialised enzymes and genetically modified crops. Several approaches have been used to determine the allergenicity of proteins in food: amino acid sequence homology to known allergens; immunologic cross-reactivity with known allergens; and resistance to degradation in gastric fluid (Kimber et al. 1999). These methods provide information about safety of new proteins, but not about the intrinsic properties of allergens that induce de novo sensitisation.

Important aspects of allergenicity are the allergen dose and route of exposure, but also the biological function of allergens. Enzymatic activity and its impact on allergenicity have been extensively debated. A large number of allergens are enzymes, such as the cysteine and serine protease from house dust mite *Dermatophagoides pteronyssinus* (*D. pteronyssinus*) Der p 1. It has been demonstrated that Der p 1 has a direct effect on a variety of immunological cell types in different in vitro systems (Schulz et al. 1995, King et al. 1998, Ghaemmaghami et al. 2001, Ghaemmaghami et al. 2002). In addition, Der p 1 has been found to break tight junctions between epithelial cells, which would facilitate its penetration through the mucosal barrier (Wan et al. 2000). The proteolytic activity has also been proven important for allergenicity in an in vivo mouse allergy model (Gough et al. 2003). Der p 1 and other enzymatic allergens have been shown to induce release of pro-inflammatory cytokines from mast cells, bronchial epithelial cells and basophils (Dudler et al. 1995, Machado et al. 1996, King et al. 1998). These and other findings suggest that enzymatic activity could explain parts of the allergenicity of some allergens, but several major allergens are not enzymes, e.g. Der p 2 and cockroach allergen Bla d 2.

### 2.3.3 Recombinant allergens

In order to study allergens at a molecular level and to examine the immunological reactions they give rise to, large amounts of pure allergens are needed. This is also essential for the development of good diagnostic and therapeutic tools (Valenta et al. 1999).

In 1988 the first allergen was cloned, Der p 1 (Chua et al. 1988). Since then, numerous allergens have been identified; the DNA that encodes the allergens has been cloned and the allergens expressed as recombinant proteins. There are several ways to identify and clone a new allergen. By using information from the DNA or protein sequence of either a homologous protein or from a short stretch of known sequence of the allergen, degenerate primers can be used to search a complementary DNA (cDNA) library (Smith et al. 1994, Saarne et al. 2003). However, by screening a cDNA expression
library, several allergens can be identified without prior sequence information (Chua et al. 1988, Shen et al. 2000). The development of the phage display approach (Smith 1985, Crameri et al. 1994) has simplified identification and cloning of allergens and is currently used in high throughput screening for new allergens (Crameri et al. 2001). Cloned allergens are usually expressed in *Escherichia coli* (*E. coli*), an expression system that is easy to handle, cost-efficient and gives good expression yields. The protein is frequently obtained in insoluble inclusion bodies, which simplifies purification but the protein often requires *in vitro* refolding. If protein activity and full IgE binding require post-translational modifications, like extensive disulphide bond formation or glycosylation, a eukaryotic expression system might be necessary. *Pichia pastoris*, Baculovirus and the mammalian cell-line Cos-1 are examples of eukaryotic expression systems that have been used successfully for production of IgE-reactive allergens, but they have some drawbacks such as hyper-glycosylation and low yield (Schmidt and Hoffman 2002).

### 2.3.4 Allergen 3D structure

Information on the 3D structure of allergens from nuclear magnetic resonance and X-ray crystallography experiments have provided us with clues of how to explain allergenicity and allergen cross-reactivity as well as giving us a base needed to study IgE epitopes. Furthermore, detailed knowledge of the 3D structure makes rational design of new forms of treatment of allergic disease possible. Today the structure coordinates of at least 26 allergens have been deposited in protein data bank (PDB), table I, together with several isoforms and mutants. Although an extensive analysis has been performed of the protein fold of characterised allergen structures, no common patterns characteristic of allergens have been found (Aalberse 2000). Nevertheless, several allergens have been observed to possess a common structural motif: a groove inside an α-β motif between an α-helix on one side and an anti-parallel β-sheet on the other side (Furmonaviciene and Shakib 2001). However, the extent to which allergenicity is determined by the protein fold is uncertain and further studies have to verify its importance. Interestingly, numerous allergens from different sources have been found to bind small hydrophobic ligands. Albumin from various mammalian sources has been shown to bind fatty acids (Curry et al. 1998) and small organic molecules. The cow, mouse and horse allergens Bos d 1, Mus m 1 and Equ c 1, respectively are lipocalins that are proposed to bind small hydrophobic molecules such as steroids and pheromones (Flower 1996). Fel d 1 belongs to the secretoglobulin family that is known to bind ligands similar to those bound by the lipocalins. The major group 2 mite allergen Der p 2 has been crystallised with two unknown ligands. Since the cavity is lined with hydrophobic and aromatic residues the ligand is proposed to be hydrophobic (Derewenda et al. 2002). The major allergen from birch, Bet v 1, is part of a group of cross-reactive allergens that belong to the Pathogenesis-related protein 10-like family. When the structure of Bet v 1 was solved, two deoxycholate molecules were found in a cavity that spans the protein (Markovic-Housley et al. 2003). The structural similarity of deoxycholate and plant steroid hormones favours the interpretation that Bet v 1 serves as a general plant steroid carrier.
Apart from the structurally characterised allergens discussed above there are other allergens that are known to bind small hydrophobic ligands, or share homology with proteins that do so.

Cross-reactivity, unlike allergenicity, is mainly dependent on allergen structure; two proteins are cross-reactive if they share epitopes, which are almost always dependent on structural features (Aalberse 2000). By comparing the structures of the *Aspergillus fumigatus* allergen manganese superoxide dismutase and its human homologue, conserved and solvent-exposed amino acids could be identified that potentially are involved in IgE mediated cross-reactivity (Flückiger *et al.* 2002). Detailed information about an IgE epitope can only be obtained from X-ray crystallography studies of an antibody in complex with an allergen. This has been achieved using Bet v 1 and a murine IgG Fab fragment, that displayed strong inhibition of allergic patients’ IgE (PDB entry 1FSK) (Mirza *et al.* 2000). The 930 Å² epitope identified was classified as discontinuous and the residue E45 was found to be an obvious candidate for mutagenesis studies aimed at destroying the IgE antibody binding properties of Bet v 1.
Table I. Allergens with available 3D structure coordinates.

<table>
<thead>
<tr>
<th>Allergen</th>
<th>Source</th>
<th>Biological function</th>
<th>PDB</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin</td>
<td>Mammals</td>
<td>Binding and/or carrier of small hydrophobic molecules</td>
<td>1AO6</td>
<td>(Curry et al. 1998, Sugio et al. 1999)</td>
</tr>
<tr>
<td>Bos d 2</td>
<td>Cow</td>
<td>Lipocalin/binding and/or carrier of small hydrophobic molecules*</td>
<td>1BJ7</td>
<td>(Rouvinen et al. 1999)</td>
</tr>
<tr>
<td></td>
<td><em>Bos domesticus</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equ c 1</td>
<td>Horse</td>
<td>Lipocalin, binding and/or carrier of small hydrophobic molecules*</td>
<td>1EW3</td>
<td>(Lascombe et al. 2000)</td>
</tr>
<tr>
<td></td>
<td><em>Equus caballus</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mus m 1</td>
<td>Mouse</td>
<td>Lipocalin/binding and/or carrier of small hydrophobic molecules</td>
<td>1MUP</td>
<td>(Böcskei et al. 1992)</td>
</tr>
<tr>
<td>Paralbumin</td>
<td>Carp</td>
<td>2 EF-hand Ca\textsuperscript{2+}-binding protein</td>
<td>1CDP</td>
<td>(Swain et al. 1989)</td>
</tr>
<tr>
<td>Api m 1</td>
<td>Honey Bee venom</td>
<td>Phospholipase A2</td>
<td>1POC</td>
<td>(Scott et al. 1990)</td>
</tr>
<tr>
<td></td>
<td><em>Apis mellifera</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Api m 2</td>
<td>Honey Bee venom</td>
<td>Hyaluronidase</td>
<td>1FCV</td>
<td>(Markovic-Housley et al. 2000)</td>
</tr>
<tr>
<td></td>
<td><em>Apis mellifera</em></td>
<td></td>
<td>1FCU</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1FCQ</td>
<td></td>
</tr>
<tr>
<td>Der p 2</td>
<td>Dust mite</td>
<td>Binding and/or transport of hydrophobic molecules*</td>
<td>1KTJ</td>
<td>(Mueller et al. 1998, Derewenda et al. 2002)</td>
</tr>
<tr>
<td></td>
<td><em>D. pteronyssinus</em></td>
<td></td>
<td>1A9V</td>
<td></td>
</tr>
<tr>
<td>Der f 2</td>
<td>Dust mite</td>
<td>Binding and/or transport of hydrophobic molecules*</td>
<td>1AHK</td>
<td>(Ichikawa et al. 1998)</td>
</tr>
<tr>
<td></td>
<td><em>D. farinae</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ves v 5</td>
<td>Yellow-jacket venom</td>
<td>Trypsin inhibitor*</td>
<td>1QNX</td>
<td>(Henriksen et al. 2001)</td>
</tr>
<tr>
<td></td>
<td><em>Vespula vulgaris</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bet v 1</td>
<td>Birch pollen</td>
<td>Homologous to PR proteins, bind and/or transport of lipid-like molecules*</td>
<td>1B6F</td>
<td>(Faber et al. 1996)</td>
</tr>
<tr>
<td></td>
<td><em>Betula verrucosa</em></td>
<td></td>
<td>1BTV</td>
<td>(Gajhede et al. 1996)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1BV1</td>
<td>(Markovic-Housley et al. 2003)</td>
</tr>
<tr>
<td>Bet v 2/</td>
<td>Birch pollen</td>
<td>actin-, PIP2-binding protein</td>
<td>1CQA</td>
<td>(Fedorov et al. 1997)</td>
</tr>
<tr>
<td>profilin</td>
<td><em>Betula verrucosa</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bet v 4</td>
<td>Birch pollen</td>
<td>2 EF-hand Ca\textsuperscript{2+}-binding protein</td>
<td>1H4B</td>
<td>(Neudecker et al. 2004)</td>
</tr>
<tr>
<td></td>
<td><em>Betula verrucosa</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hev b 6</td>
<td>Latex</td>
<td>Unknown</td>
<td>1HEV</td>
<td>(Rodriguez-Romero et al. 1991)</td>
</tr>
<tr>
<td></td>
<td><em>Hevea brasiliensis</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hev b 8/</td>
<td>Latex</td>
<td>Actin-, PIP2-binding protein *</td>
<td>1G5U</td>
<td>-</td>
</tr>
<tr>
<td>profilin</td>
<td><em>Hevea brasiliensis</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pru Av 1</td>
<td>Cherry tree pollen</td>
<td>Homologous to PR proteins, carrier of hydrophobic molecules *</td>
<td>1E09</td>
<td>(Neudecker et al. 2001)</td>
</tr>
<tr>
<td></td>
<td><em>Prunus avium</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* not verified
### Table I. Allergens with available 3D structure coordinates, continued.

<table>
<thead>
<tr>
<th>Allergen</th>
<th>Source</th>
<th>Biological function</th>
<th>PDB</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amb t 5</td>
<td>Ragweed pollen</td>
<td>Unknown</td>
<td>1BBG</td>
<td>(Metzler et al. 1992)</td>
</tr>
<tr>
<td>Ara t 8/</td>
<td>Arabidopsis thaliana</td>
<td>Actin-, PIP2-binding protein *</td>
<td>1A0K</td>
<td>(Thorn et al. 1997)</td>
</tr>
<tr>
<td>Phl p 1</td>
<td>Timothy grass pollen</td>
<td>β-expansin, papain-related proteinase</td>
<td>1N10</td>
<td></td>
</tr>
<tr>
<td>Phl p 2</td>
<td>Timothy grass pollen</td>
<td>Unknown</td>
<td>1WHO</td>
<td>(De Marino et al. 1999)</td>
</tr>
<tr>
<td>Phl p 5</td>
<td>Timothy grass pollen</td>
<td>RNAs</td>
<td>1L3P</td>
<td>(Rajashankar et al. 2002)</td>
</tr>
<tr>
<td>Phl p 6</td>
<td>Timothy grass pollen</td>
<td>Unknown</td>
<td>1NLX</td>
<td></td>
</tr>
<tr>
<td>Phl p 7</td>
<td>Timothy grass pollen</td>
<td>2 EF-hand Ca²⁺-binding protein</td>
<td>1K9U</td>
<td>(Verdino et al. 2002)</td>
</tr>
<tr>
<td>Llpr 10.1</td>
<td>Yellow Lupin</td>
<td>Homologous to PR proteins,</td>
<td>1ICX</td>
<td>(Biesiadka et al. 2002)</td>
</tr>
<tr>
<td>Gly m 1</td>
<td>Soybean</td>
<td>Unknown</td>
<td>1HYP</td>
<td>(Baud et al. 1993)</td>
</tr>
<tr>
<td>Zea m 14</td>
<td>Maize</td>
<td>Lipid-transfer protein</td>
<td>1MZL</td>
<td>(Shin et al. 1995)</td>
</tr>
<tr>
<td>Asp f 6</td>
<td>Aspergillus fumigatus</td>
<td>Oxidoreductase</td>
<td>1KKC</td>
<td>(Flückiger et al. 2002)</td>
</tr>
</tbody>
</table>

* not verified

### 2.4 MITE ALLERGY

#### 2.4.1 Sensitisation to mites and mite allergens

Mites are found almost all over the world. Mites and ticks belong to the group of arthropods and to the subclass Acari of the class Arachnida, and more than 30 000 mite and tick species have been identified, Fig. 2. Some live on plants (leaves and fruits) or animals (skin and bird feathers), while others aid in recycling of nutrients by feeding on dead material from plants and animals. Several mites are predators, that feed on other mites and insects (Fernandez-Caldas 2002).

The dust mites *D. pteronyssinus* and *D. farinae* are found in homes in humid regions all around the world (Arlian et al. 2002) and are two of the dominating causes of allergic sensitisation (Platts-Mills et al. 1997). In Sweden, Plaschke et al. have shown that 7% - 14% of randomly selected subjects show positive reactions to *D. pteronyssinus* in skin prick tests (Plaschke et al. 1996). Some other mite species that are sources of mite sensitisation are *Euroglyphus maynei* (*E. maynei*), which is common in houses in humid regions like the southern part of the United States (Arlian et al. 1992), and *Blomia tropicalis* (*B. tropicalis*) which is found in the tropical and the sub-tropical parts of the world (Montealegre et al. 1997, Mariana et al. 2000). The mite *L. destructor* is, besides *D. pteronyssinus* and *D. farinae*, one of the most common mite species.
*L. destructor* is a storage mite found both in rural (Boström et al. 1997, Franz et al. 1997) and urban (Warner et al. 1999) environments and has been shown to cause sensitisation and allergic disease (Cuthbert et al. 1979, van Hage-Hamsten et al. 1985, van Hage-Hamsten et al. 1988, Ebner et al. 1994, Kronqvist et al. 1999, Müsken et al. 2000, Vidal et al. 2004). Spider mites are a group of mites that cause the human double trouble. They are pests on various crops grown in green houses and fruit orchards and may cause sensitisation among green house workers (Delgado et al. 1997, Johansson et al. 2003). To combat these mites and other pests, predatory mites, such as *Hypoaspis miles*, *Phytoseiulus persimilis* and *Amblyseius cucumeris*, are cultured for commercial use in biological pest control. The predatory mites are added onto the plants where they actively search for their prey and destroy them. Recently, however, these mites have also been shown to induce allergic sensitisation in green house workers (Groenewoud et al. 2002, Johansson et al. 2003).

![Figure 2. Classification of some common mites.](image)

Mites are one of the most well-characterised allergen sources, with a total of 43 registered allergens in the IUIS allergen database. More than 30 components from the *Dermatophagoides* spp. alone (Baldo et al. 1989) and 21 from *L. destructor* have been shown to bind IgE (Johansson et al. 1991). Mite allergens are divided into 19 groups of proteins; these proteins share sequence homology and have similar molecular weights within the group. The different allergen groups have various biological functions and a range of IgE reactivity (Thomas et al. 2002).

The two most well studied mite allergen groups are groups 1 and 2:

The group 1 allergens are 25- kDa glycoproteins and display homology to cysteine and serine proteases. Their enzymatic activity and the effect on allergenicity have been discussed above and in detail elsewhere (Thomas et al. 2002). Group 1 allergens have been identified in the species *B. tropicalis*, *D. farinae*, *Dermatophagoides microceras*, *D. pteronyssinus* and *Dermatophagoides siboney* (*D. siboney*). Der p 1 has been shown
to be recognised by up to 100% of *D. pteronyssinus* allergic patients and 50-70% of the IgE antibodies against a *D. pteronyssinus* extract were directed to Der p 1 (Thomas *et al.* 2002).

The group 2 allergens have a molecular mass of approximately 15- kDa and display among the highest frequencies of IgE reactivity of all mite allergens. The group 2 allergens of *Dermatophagoides* spp. bind IgE in up to 80% of patients sensitised to *D. pteronyssinus* and *D. farinae* (Thomas *et al.* 2002). Lep d 2, the dominating allergen in *L. destructor*, is recognised by over 70% of *L. destructor* sensitised patients in vitro (Johansson *et al.* 1999) and 60% in skin prick test (Kronqvist *et al.* 2000). The biological function of the group 2 allergens is currently not known, but they are proposed to be binding hydrophobic ligands (Derewenda *et al.* 2002). Group 2 allergens have been identified in *D. farinae*, *D. pteronyssinus*, *D. siboney*, *E. maynei*, *Glycyphagus domesticus*, *L. destructor* and *Tyrophagus putrescentiae*.

The major allergen Lep d 2 has been found in two isoforms, Lep d 2.01 and Lep d 2.02, differing in 13 amino acids and numerous nucleotides (Schmidt *et al.* 1995). Isoform Lep d 2.01 was found as two variants with identical amino acid sequence, differing only at the DNA level. In accordance with WHO/IUS allergen nomenclature (WHO/IUS 1995), these were named Lep d 2.0101a, formerly Lep d 2.0101, and Lep d 2.0101b, formerly Lep d 2.0102. Lep d 2.02 was found only as one variant and named Lep d 2.0201. Skin prick testing of 41 *L. destructor* sensitised individuals displayed similar reactivity to Lep d 2.01 and Lep d 2.02 (Kronqvist *et al.* 2000).

The presence of different allergen isoforms and variants is well known, and that needs to be taken into account when recombinant allergens are used in diagnostics and immunotherapy. In addition, different patterns of polymorphisms have been found in wild mites and mites cultured for diagnostic and therapeutic purposes (Chua *et al.* 1996, Smith *et al.* 2001). The importance of polymorphisms has been thoroughly investigated for several *D. pteronyssinus* and *D. farinae* allergens (Thomas *et al.* 1992, Chua *et al.* 1996, Yuuki *et al.* 1997, Smith *et al.* 2001). In these studies, group 2 allergen genomic and cDNA sequences from cultured mites have been shown to contain mutations, resulting in three to five amino acid substitutions and several silent mutations. Some of these mutations are also observed in mites taken straight from their natural environment. Compared to Der p 2, sequences from Der p 1 clones show fewer base pair substitutions, but the base pair changes of Der p 1 more often lead to amino acid changes (Smith *et al.* 2001). Furthermore, it has been shown that a single amino acid substitution can have a major effect on T-cell responses. A recent study revealed that peptides representing various Der p 1 T-cell epitopes containing polymorphic residues differed in their ability to induce T-cell proliferation (Smith *et al.* 2001). In addition, in the same study, no difference was found in the ability of four variants of Der p 2 to stimulate proliferation of peripheral blood mononuclear cells (PBMC). For *L. destructor* it has been shown that the Lep d 2.0101 isoform induces stronger T-cell responses than Lep d 2.0201 (Eriksson *et al.* 2001).
2.5 CAT ALLERGY

2.5.1 Sensitisation to cat and cat allergens

Sensitisation to the domestic cat (*Felis domesticus*) is a common cause of allergic disease, with a prevalence of approximately 10% in the westernised world (Freidhoff *et al.* 1984, Roost *et al.* 1999). Cat allergens are abundant in society, even in environments where cats are not present, such as schools and public transportation (Partti-Pellinen *et al.* 2000, Almqvist *et al.* 2001), and are therefore difficult to avoid. In Sweden, cat is one of the most common causes of sensitisation and 15% of randomly selected populations were found positive to cat in skin prick test (Plaschke *et al.* 1996).

A dose-response relationship between exposure to cat and sensitisation has been shown (Wahn *et al.* 1997). However, during the past few years results have been published showing that some subjects who live with cats and are exposed to very high levels of cat allergens can develop a specific form of tolerance. These subjects display a modified T\(_2\) response to cats with high titres of allergen-specific IgG and IgG4 antibodies without IgE, a phenomenon that has not been observed for other allergens (Custovic *et al.* 2001, Platts-Mills *et al.* 2001). Furthermore, this type of response was not associated with an increased risk of allergic symptoms or asthma (Perzanowski *et al.* 2002).

The most prominent and potent allergen in cat dander is termed Fel d 1 (originally called Cat allergen 1) and was identified more than three decades ago (Ohman *et al.* 1974). Fel d 1 elicits IgE responses in 90-95% of patients with cat allergy (van Ree *et al.* 1999) and accounts for 60-90% of the total allergenic activity in cat dander (Kleine-Tebbe *et al.* 1993). The allergen has been found in cat saliva and pelt, and in salivary, lachrymal and sebaceous glands as well as in squamous epithelial cells (Anderson *et al.* 1985, Bartholome *et al.* 1985, van Milligen *et al.* 1990, Charpin *et al.* 1991). Fel d 1 is a 35- kDa tetrameric glycoprotein (Kristensen *et al.* 1997) formed by two heterodimers (Duffort *et al.* 1991). Each dimer is composed of two chains derived from independent genes, chain 1 comprising 70 residues and chain 2, 90 or 92 residues (Morgenstern *et al.* 1991, Griffith *et al.* 1992). The two isoforms of the second chain are expressed in the skin and the saliva, respectively (Griffith *et al.* 1992). The chains are linked by three disulphide bonds, which are formed between Cys3 of chain 1 and Cys73 of chain 2, Cys44 of chain 1 and Cys48 of chain 2, and between Cys70 of chain 1 and Cys7 of chain 2 (Bond *et al.* 1993, Kristensen *et al.* 1997). Several attempts have been made to characterise both T- and B-cell epitopes of Fel d 1. The immunodominant T-cell epitopes identified have been localised mainly on chain 1 (Counsell *et al.* 1996, Mark *et al.* 1996). However, in a recent study the N termini of both chains were shown to stimulate strong T-cell proliferation (Reefer *et al.* 2004). To study IgE binding epitopes, overlapping peptides have been used. Three peptides were found to be recognised more often than the others; however, only sera with high levels of IgE antibody to cat detected the peptides (van Milligen *et al.* 1994).

No biological function has yet been determined for Fel d 1. However, orthologues of Fel d 1, the secretoglobins (Klug *et al.* 2000), have been described as anti-inflammatory cytokine-like molecules capable of binding various hydrophobic molecules (Mukherjee *et al.* 1999). The human orthologue, Clara cell 10- kDa protein, shares 22% sequence
Molecular characterisation of major allergens from mite and cat

identity with chain 1 of Fel d 1. In addition, Fel d 1 has been associated with gelatin- and fibronectin-degrading activity. However, no serine peptidase catalytic triad has been found in Fel d 1, which indicates that the activity is due to non-catalytic site interactions or that an active contaminant was co-purified with Fel d 1 (Ring et al. 2000).

A recombinant Fel d 1 molecule with properties identical to the natural protein has been urgently needed for many years. Several attempts have been made to refold recombinant Fel d 1 into a native-like allergen but with only partial success (Bond et al. 1993, Keating et al. 1995, Slunt et al. 1995, Vailes et al. 2002). Although, a mixture of the separate chains has been proven useful for in vitro allergy diagnostics (Slunt et al. 1995, van Ree et al. 1999). Nevertheless, a soluble and correctly folded recombinant molecule would be of value, not only for diagnosis and treatment, but also for molecular studies of the Fel d 1 protein. Recently, the appropriate conditions for expression and in vitro folding of recombinant Fel d 1 have been reported. By using a direct fusion of the Fel d 1 chains, two constructs with properties similar to the natural allergen were generated, Fel d 1 (1+2) and Fel d 1 (2+1), Fig. 3 (Grönlund et al. 2003). Fel d 1 (2+1) will hereafter be referred to as recombinant (r)Fel d 1. Indeed, linking of the two chains seems to have provided the physical contacts needed to create in vitro the correct disulphide-bonding pattern and protein fold. Moreover, recombinant Fel d 1 displays a secondary structure comparable to that of the natural protein, as revealed by circular dichroism, and it forms a dimer corresponding to the natural tetramer. Most importantly, the in vitro immunoreactivity of the stable rFel d 1 is indistinguishable from that of the natural allergen (Grönlund et al. 2003).

Figure 3. Expression constructs used for Fel d 1 production.

A, Fel d 1 (1+2). B, Fel d 1 (2+1). Chain 1, 2 and the six histidine tags are coloured in white, grey and black, respectively.

Cat albumin or Fel d 2 is a minor allergen from cat and is recognised by IgE from 35% of cat-allergic patients (Hilger et al. 1996, Reininger et al. 2003). Albumin is a cross-reactive allergen and IgE from albumin-sensitised subjects recognise albumin from various animals (Spitzauer et al. 1995).

Fel d 3 or cat cystatin is another minor cat allergen, it is recognised by at least 10% of cat-allergic patients. Sequence comparisons and homology modelling have revealed that Fel d 3 contains the conserved cysteine protease inhibitor signature and two of three lipocalin motifs (Ichikawa et al. 2001).
Recently four new cat allergens have been added to the IUIS allergen database. Fel d 4, which belongs to the lipocalin family, and Fel d 5, 6 and 7 that are cat immunoglobulins A, M and G, respectively. Cat IgA and IgM are recognised by 60-70% and cat IgG by 5% of cat-sensitised subjects (Adedoyin et al. submitted). No data on IgE reactivity for Fel d 4 has been reported.

2.6 DIAGNOSIS AND TREATMENT OF ALLERGIC PATIENTS

2.6.1 Current diagnostics and therapy

Currently two techniques are routinely used to diagnose allergic sensitisation: in vivo skin prick test and in vitro measurement of allergen-specific IgE in serum. In skin prick tests allergens are introduced into the skin. This leads to mast cell degranulation and a weal and flare reaction in sensitised individuals. The test is evaluated based on the size of the weal. Circulating IgE can be measured in vitro with various commercial systems. In this thesis, the Pharmacia CAP System™ was used to measure allergen-specific IgE. The CAP system is calibrated against the international reference standard for human IgE. The level of IgE is divided into six radio allergosorbent test (RAST) classes: class one contains 0.35-0.7 kU/L; class two 0.7-3.5 kU/L; class three 3.5-17.5 kU/L; class four 17.5-50 kU/L; class five 50-100 kU/L and class six >100 kU/L of allergen-specific IgE.

To treat the allergic symptoms, different pharmacological substances are commonly used. However, the effects of the drugs are transient. The only treatment for allergic disease that provides a long-lasting relief of symptoms is allergen-specific immunotherapy (SIT), in which gradually increasing doses of allergens are administered to the patient in order to induce unresponsiveness (Bousquet et al. 1998). Although the clinical efficacy of SIT has been documented, the procedure entails a risk of local and systemic side effects (Bousquet et al. 1998, Valenta 2002). The immunological mechanisms behind the induced allergen unresponsiveness are not fully understood, but the induction of an IgG antibody response capable of inhibiting IgE-binding to the allergen (blocking IgG antibodies) has been proposed as an important factor for successful outcome in SIT (Ball et al. 1999, Mothes et al. 2003). However, blocking IgG antibodies have not always been associated with a successful clinical outcome (Djurup and Malling 1987, Birkner et al. 1990). The importance of different T-cell subsets during SIT has also been investigated using PBMC and T-cell clones. The levels of IL-4 and IL-5 were shown to decrease after SIT (Jutel et al. 1995, Ebner et al. 1997), and that of IFN-γ has been shown to increase (Jutel et al. 1995). The changes in cytokine secretion would thereby shift the cytokine production from a Th2 type (mainly IL-4, IL-5 and IL-13) to a Th1 type (mainly IFN-γ). The role of IL-10 and transforming growth factor β producing regulatory T-cells as modulators of the immune response is not clear, but an increase in IL-10 secreted by allergen stimulated PBMC from SIT treated patients has been found (Francis et al. 2003, Jutel et al. 2003, Nouri-Aria et al. 2004). Furthermore, IL-10 has been shown to decrease IL-4 induced IgE transcription by PBMC and increase IL-4 induced IgG4 transcription (Jeannin et al. 1998). Local expression of IL-10 has also been detected in the nasal mucosa of grass pollen allergic patients after SIT and during the following pollen season. The induction
of IL-10 in the mucosa was found to be linked to increased serum levels of allergen-specific IgG and IgG4 (Nouri-Aria et al. 2004).

2.6.2 Future diagnostics and therapy

The allergens used today in both diagnostics and therapy are based on aqueous allergen extracts. However, allergen extracts prepared from natural sources have several disadvantages. They contain a number of different allergens, as well as carbohydrates, histamine (Williams et al. 1992) and endotoxins (Trivedi et al. 2003) that may interfere in allergy testing. Moreover, the extract may be contaminated with allergens from other sources (van der Veen et al. 1996) and the allergen concentration may vary (Rosenbaum et al. 1996). The batch-to-batch variations in allergen extracts create problems with diagnostics and therapy and complicate comparisons between different studies. Another potentially important problem related to crude allergen extracts is that new, therapy induced IgE-reactivities to extract components can arise (Ball et al. 1999, Moverare et al. 2002). The availability of highly pure and well defined recombinant allergens not only overcomes these problems but also permits component-resolved diagnostics (CRD) (Valenta et al. 1999). In CRD, each specific allergen that a patient is sensitised to is identified. Recently, a microarray with a large number of allergens spotted onto glass plates has been developed (Hiller et al. 2002). The allergen chips will in the future allow fast determination of a patient’s allergen reactivity profile, and will require only a small amount of serum.

The production of recombinant allergens has also opened new opportunities for refined treatment. First, it will be possible to treat a patient with only the allergen he or she is sensitised against instead of using whole extracts. Second, recombinant allergens can be modified into hypoallergens, which have reduced allergenic activity but retained T-cell reactivity. The application of hypoallergens in SIT could lead to reduced risk of side effects (Akdis and Blaser 2000, Valenta 2002), making it possible to administer higher doses and use fewer injections. Several types of hypoallergenic derivatives have been created (Akdis and Blaser 2000). Among these are chemical modification of allergens, like conjugation of allergens with polyethylene glycol (Lee and Sehon 1977). However, the progress in DNA technology has now made it possible to tailor-make hypoallergens. One example is the Bet v 1 hypoallergens, which have been extensively characterised. With polymerase chain reaction (PCR) based construction, the allergenicity of Bet v 1 has been almost abolished by dividing the allergen into two fragments (Vrtala et al. 1997, Vrtala et al. 1999). Unfortunately, when this approach is used important T-cell epitopes may also be removed. This has been avoided by using another method, oligomerisation, where three copies of Bet v 1 are expressed in sequence to form a Bet v 1 trimer (Vrtala et al. 2001). Both the Bet v 1 trimer and the Bet v 1 fragment mix display considerably decreased IgE binding in vivo (van Hage-Hamsten et al. 1999, Pauli et al. 2000) and reduction of allergic symptoms in nasal provocation studies (van Hage-Hamsten et al. 2002). Another approach to generate hypoallergens is based on site-directed mutagenesis of residues crucial for the protein fold and IgE-binding reactivity. In the group 2 mite allergens this has been achieved by changing cysteines involved in disulphide bridges to serines (Smith and Chapman 1996, Takai et al. 1997, Olsson et al. 1998). The serine mutants displayed a marked
reduction in IgE binding capacity (Smith and Chapman 1996, Takai et al. 1997, Olsson et al. 1998, Kusunoki et al. 2000, Kronqvist et al. 2001). In addition, it has been shown in mouse that by injecting only T-cell epitope-containing peptides, allergen unresponsiveness can be induced (Briner et al. 1993, Hoyne et al. 1993). However, when this approach has been applied on humans allergic to bee venom and cat the results have been contradictory (Norman et al. 1996, Simons et al. 1996, Müller et al. 1998, Pene et al. 1998, Oldfield et al. 2002).
3 AIMS OF THE THESIS

The purpose of this thesis was to characterise allergens from two of the most important sources of allergens worldwide: cats and mites.

The specific aims of the individual papers are:

I. To investigate the occurrence and importance of polymorphisms in Lep d 2, a major allergen from *Lepidoglyphus destructor*, derived from different sources.

II. To establish a system for purification, folding and structural studies of recombinant Fel d 1, the major cat allergen.

III. To produce seleno-methionine substituted Fel d 1 and determine the 3D structure of Fel d 1.

IV. To test a new approach for rational design of hypoallergens for immunotherapy, using structural information and knowledge of B- and T-cell epitopes of an allergen.
4 MATERIAL AND METHODS

This section is an overview of the methods used in this thesis. A more detailed description of the methods is given in the individual papers.

4.1 MITES

(I) Cultured *L. destructor* mites were purchased from Allergon AB, Ängelholm, Sweden. Wild *L. destructor* mites were identified and isolated from a hay sample collected from a farm on the Swedish island Gotland.

4.2 SERUM SAMPLES

(I) Sera from six *L. destructor* allergic farmers from Gotland were used in immunoblotting and enzyme-linked immunosorbent assay (ELISA) inhibition experiments. The sera were RAST positive to *L. destructor* extract and CAP positive to recombinant Lep d 2.0101 (Johansson et al. 1999). One serum sample with a negative RAST to *L. destructor* was used as a negative control.

(IV) A serum pool derived from 31 individuals with positive CAP values to cat was used in ELISA inhibition experiments. For T-cell proliferation experiments blood was obtained from nine cat-allergic patients. All patients were CAP positive to cat dander and positive to rFel d 1 in ELISA. Blood from four of these patients was used in the basophil activation assay. None of the patients had used antihistamines or inhalant corticosteroids for at least 5 days prior to the blood donation, and none had previously received immunotherapy. For control purposes, blood was obtained from two non-atopic individuals.

4.3 PCR AMPLIFICATION AND CLONING

(I) Genomic DNA encoding Lep d 2 was amplified using one mite as DNA template. The PCR products were cloned into pCR4-TOPO vectors, and the plasmids were transformed into *E. coli* TOP10.

(IV) Short sequences encoding parts of known T-cell epitopes in Fel d 1 were duplicated by PCR. The PCR products containing different Fel d 1 constructs were ligated into pET-20b vectors.

All DNA sequencing was carried out on an ABI 377 Sequencer. Primers that were designed to sequence both strands of the PCR products were used.

4.4 SITE DIRECTED MUTAGENESIS

(I, IV) The site directed mutagenesis experiments were designed to change one codon at a time in Lep d 2 and rFel d 1. The nucleotide exchanges were confirmed by DNA sequencing, and plasmids containing the correct sequences were transformed into *E. coli* BL21 (DE3) pLysS for expression.
4.5 GENOMIC DNA EXTRACTION, SOUTHERN BLOT ANALYSIS

(I) High MW genomic DNA was extracted from cultured *L. destructor* mites. Five micrograms of DNA was digested separately with the following restriction enzymes: *Eco RI*, *Tse I*, *Bsa I* and *Ban I*. Southern blot and hybridisation were performed according to standard protocols before autoradiography. DNA encoding Lep d 2 was amplified by PCR from one cultured mite and used as a hybridisation probe in Southern blot analysis. The PCR product was labelled with $^{32}$P-dCTP.

4.6 PROTEIN EXPRESSION AND PURIFICATION

(I, II, III, IV) Recombinant proteins were expressed as C-terminal tagged hexahistidine fusion proteins using the pET-expression system and purified using metal chelate affinity chromatography.

(II, III, IV) Further purification using size exclusion and (II, III) ion exchange chromatography was carried out to increase the purity of Fel d 1.

(III) Expression of seleno-methionine (SeMet)-substituted Fel d 1 was performed by inhibiting the methionine biosynthesis pathway by adding high concentrations of isoleucine, lysine, phenylalanine, valine, threonine and SeMet.

4.7 ELISA

(I, IV) In IgE ELISA inhibition experiments, patients’ sera were pre-incubated with the Lep d 2 or Fel d 1 protein in serial dilutions before being added to 96-well ELISA plates. The wells were coated with the Lep d 2 variants or rFel d 1. Detection was carried out with rabbit anti-human IgE and alkaline phosphatase conjugated goat anti-rabbit IgG. Finally, incubation with substrate (p-nitrophenyl phosphate disodium) was performed in the dark; thereafter the absorbance was measured at 405 nm. Inhibition values were calculated by using the following formula: % inhibition = 100 - $100(A/B)$, where A is the absorbance value obtained for a serum incubated with allergen and B is the value for the same serum incubated with diluent.

(I) The cytokines IFN-γ and IL-5 were measured with commercial sandwich ELISA kits.

4.8 PBMC STIMULATION

(I) Cytokine secretion was assayed in long-term PBMC cultures. The cells were stimulated with the Lep d 2 proteins (5 µg·mL$^{-1}$) at day zero and were further stimulated with IL-2 at day five and eight of culture. Supernatants were collected at day eleven and kept at −20 °C until assessment by ELISA.

(IV) The lymphoproliferative activity of rFel d 1 and the Fel d 1 derivatives (2 and 6 µg·mL$^{-1}$) were assayed in triplicates after incorporation of $[^3]$H]thymidine by PBMC in seven-day cultures. For control purpose PBMC were stimulated with tetanus toxoid and tuberculin protein derivate or medium alone. Stimulation index (SI) was calculated as mean cpm for antigen-stimulated PBMC divided by mean cpm for PBMC cultured in medium alone. The SIs presented are the highest values obtained, regardless of the allergen concentration used.

(I, IV) All allergens were purified from contaminating endotoxins before being added to the cell cultures.
4.9 BASOPHIL ACTIVATION

(IV) To assay activation and degranulation of basophils, blood samples were incubated with rFel d 1 or one of the Fel d 1 derivatives. The basophils were then labelled with phycoerythrin-conjugated anti-CD203c and fluorescein-isothiocyanate conjugated anti-CD63 monoclonal antibody (mAb) and assessed by two-colour flow cytometry. The basophil marker CD203c was used to set a basophil gate including only CD203c⁺ cells. Two hundred events were collected within the gate and were analysed for CD63-positivity (marker for granule release). The magnitude of basophil activation was calculated as the percentage of CD63-positive events among the 200 events in the basophil gate (CD63⁺/CD203c⁺).

4.10 FEL D 1 CRYSTALLISATION

(II, III) The optimal conditions of crystal growth were found by testing different conditions such as temperature, pH, and concentration of precipitant/buffer/protein and different detergents.

4.11 DATA COLLECTION AND PROCESSING

(II, III) To improve the data collected, the crystals were soaked in a cryoprotectant solution (20% of 2-methyl-2,4-pentanediol (MPD)) before flash freezing in a cold nitrogen stream. Two data set were collected, one at beamline I711 at MAX-II, Lund, Sweden (II), and the other at ID29 beam line, ESRF, Grenoble, France (III). Data were processed using MOSFLM, scaling and reduction of the data were performed with CCP4 programs. The space group and cell dimensions were determined using the auto-indexing option of MOSFLM and by the analysis of pseudo-precession images.

4.12 PHASING, MODEL BUILDING AND REFINEMENT

(III) The localisation of the positions of the anomalous scatters and the single anomalous diffraction phasing was carried out using the program SOLVE. The phases were extended to 2.2 Å using non-crystallographic symmetry averaging and solvent flattening using the program RESOLVE. This resulted in an interpretable electron density map. Automated chain tracing by the program RESOLVE gave an initial model and the remaining parts were built with the program O. The native data set to 1.85 Å resolution was used for the refinement, which was carried out using the program REFMAC5. Isotropic, individual B-factor refinement was used throughout the procedure. The final validation included calculation of a composite omit map and analysis using PROCHECK. Structural comparisons were carried out with O using default parameters.
5 RESULTS AND DISCUSSION

5.1 LEP D 2 POLYMORPHISMS IN LEPIDOGlyphUS DESTRUCTOR MITES (I)

The objective of this study was to describe the genomic organisation and polymorphisms of the major allergen Lep d 2 in wild and cultured *L. destructor* mites. In addition, the effects of polymorphisms on B- and T-cell epitopes were evaluated in inhibition ELISA and T-cell cytokine release assays.

Eight cultured mites from a commercial source and ten wild mites from a hay dust sample were used as templates to amplify genomic Lep d 2 DNA. Two PCR products, 400 bp and 480 bp in size, were obtained from each PCR reaction using both wild and cultured mites as template. The 400 bp band corresponds well to the size of the known Lep d 2 cDNA (Schmidt et al. 1995). Both PCR products were subsequently ligated into a cloning vector and sequenced. Sequence analysis revealed that the 400 bp PCR products corresponded to the known Lep d 2 cDNA sequence with a few exceptions. Sequence analysis of clones containing the 480 bp PCR products revealed the presence of 76 or 75 nucleotides interrupting the coding sequence of both Lep d 2.0101 and Lep d 2.0201, respectively. The sequence was inserted after base pair 73 in the cDNA-sequence (Schmidt et al. 1995) corresponding to amino acid nine in the mature protein, Fig. 4. The inserted sequence most likely corresponds to an intron since it begins with GT and ends with AG, which is in agreement with 5´and 3´splice junctions surrounding intron sequences (Mount 1982). The size of the intron correlates well with introns previously reported in Der p 2 (80-83 bp) (Chua et al. 1996) and Der f 2 (87 bp) (Yuuki et al. 1997).

![Figure 4. Amino acid sequence alignment of Lep d 2 variants. Identities with Lep d 2.0101 are indicated with full stops. The amino acids changed using site directed mutagenesis are indicated in bold. An arrowhead marks the site of the intron.](image-url)
Analysis of sequences from six to eight clones originating from each of the ten wild mites revealed that the clones are clustered into two groups with a high degree of sequence identity with each other. We found clones from all ten mites that were identical or similar to Lep d 2.0101, and in all but one mite to Lep d 2.0201. Fifty percent of the clones contained nucleotide changes resulting in one to three amino acid substitutions compared to the known Lep d 2 isoforms. The two most consistent variations were in the nucleotides encoding the amino acids at positions 55 and 102 in the mature protein of Lep d 2.0101 and Lep d 2.0201, respectively, Fig. 4. Since these substitutions differed only in one amino acid compared to known isoforms, they were considered to be variants of the Lep d 2.01 and Lep d 2.02 isoforms and were designated Lep d 2.0102 and Lep d 2.0202, respectively, according to the nomenclature (WHO/IUS 1995). Figure 5 displays the relationship between the different variants.

Figure 5. Lep d 2 isoforms and variants. Lep d 2 is present as two distinct isoforms, Lep d 2.01 and Lep d 2.02, differing in 13 amino acids. The variants of each isoform differ only in a few amino acids with the exception of Lep d 2.0101a and Lep d 2.0101b, which have identical amino acid sequences but differ on the DNA level.

More than one variant was often identified from a single mite. For the Lep d 2.01 isoform, variant Lep d 2.0101 was found in all the ten mites while Lep d 2.0102 was found in seven out of ten mites analysed. The distribution of isoform Lep 2.02 variants showed that Lep d 2.0201 was present in nine out of the ten mites and Lep d 2.0202 in six, table II. In addition to these variants with amino acid substitutions, several silent mutations were found in almost all sequences investigated.

Sequences from six to eight clones originating from each of the eight cultured mites resemble those from wild mites regarding the distribution of Lep d 2.0101 and Lep d 2.0201-like clones. The same patterns of base pair changes were seen in the cultured mites as in the wild mites. However, the frequencies of the amino acid substitutions were different. The variants Lep d 2.0101 and Lep d 2.0102 were both found in seven out of eight mites. Lep d 2.0201 was found in all eight mites and Lep d 2.0202 in one, table II.

Table II. Frequencies of Lep d 2 variants.

<table>
<thead>
<tr>
<th></th>
<th>Lep d 2.0101</th>
<th>Lep d 2.0102</th>
<th>Lep d 2.0201</th>
<th>Lep d 2.0202</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild L. destructor (n=10)</td>
<td>10</td>
<td>7</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Cultured L. destructor (n=10)</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>

Sequencing of the Lep d 2 gene revealed several silent mutations but only a few mutations resulted in amino acid substitutions. This result is in analogy with data obtained from sequence analyses of Der p 2 (Smith et al. 2001). Furthermore, all
analysed clones were highly homologous to the Lep d 2 cDNA sequences published earlier (Schmidt et al. 1995), i.e. to Lep d 2.0101 or Lep d 2.0201, which indicates an evolutionary divergence of two sequences corresponding to two main isoforms, in analogy with what has been found earlier for Der p 2 (Smith et al. 2001). The commercially available mites used in this study are cultured mites that have been grown isolated for years without introduction of new mites from other sources (personal communication, A. Anderson, Allergon AB). This could explain the disparate frequency of the variants found in cultured and wild mites.

Data from sequencing, showing the presence of Lep d 2 with and without an intron, indicate that multiple copies of the Lep d 2 gene are present in the *L. destructor* genome. To investigate this possibility, Southern blot analysis was performed on genomic *L. destructor* DNA. The DNA was digested with each of four enzymes: *Eco RI* recognising no restriction site, *Tse I* and *Bsa I* recognising single and *Ban I* recognising two restriction sites in the Lep d 2 complete DNA sequence. Hybridisation of the Lep d 2 probe with *Eco RI*-cleaved DNA gave rise to two bands, and at least three bands were seen when the DNA was cleaved using *Tse I*, *Bsa I* or *Ban I*, Fig. 6. These data support the idea that there is more than one copy of the Lep d 2 gene at different loci in the *L. destructor* genome. These results are in contrast to those obtained for Der f 2, where only a single Der f 2 gene is found in the genome (Yuuki et al. 1997). Hence, the substitutions found in the Lep d 2 gene are probably not only due to polymorphisms within the Lep d 2 gene, but also to multiple copies of the gene in individual mites.

**Figure 6.** Southern blot analysis of Lep d 2.

Genomic *L. destructor* DNA digested with each of four enzymes: *Eco RI* recognising no restriction site, *Tse I* and *Bsa I* recognising single and *Ban I* recognising two restriction sites in the Lep d 2 complete DNA. Lane 1, *Bsa I*; lane 2, *Ban I*; lane 3, *Eco RI*; lane 4, *Tse I*.

ELISA inhibition experiments were performed to investigate if the Lep d 2 variants found in cultured and wild mites have different IgE binding properties. A dose-dependent inhibition was observed with all variants in six different sera. The IgE binding capacity for the two isoforms was evaluated separately. In all experiments 100% inhibition was reached. Figure 7 shows the inhibition results obtained with two of the six sera using the two variants of the isoform Lep d 2.01. For both sera, lower concentration of the variant Lep d 2.0102 was needed to reach 50% inhibition compared to the variant Lep d 2.0101 regardless of whether homologous or heterologous inhibition was performed. Similar results were obtained with the isoform
Lip d 2.02, where a lower concentration of the variant Lip d 2.0202 was needed to reach 50% inhibition compared to the Lip d 2.0201 variant, although the difference was less pronounced. The same pattern could be seen using all sera, table III Studies of antibody epitopes in the orthologue Der p 2, have shown that residues at position 55 and 102 are within the predicted B-cell epitopes and could be important in IgE binding (Mueller et al. 2001). On the other hand, the crystal structure of Der p 2 has been used in homology modelling of Lip d 2.0101 (Gafvelin et al. 2001) and according to this model, amino acids 55 and 102 are buried inside the protein core of Lip d 2 and not exposed on the surface of the molecule (personal communication, D. Benjamin). Thus, the difference in IgE binding is probably not caused by direct antibody interaction of the side chains of the variable amino acid residues but rather by more subtle changes in the tertiary structure.

**Figure 7.** ELISA inhibition.

ELISA inhibition of IgE-binding to Lip d 2.0101 (a and c), and Lip d 2.0102 (b and d) on solid phase. Inhibition curves obtained with Lip d 2.0101 (— — — —), Lip d 2.0102 (—— — —) as indicated in the figure. Serum no. 6 was used in a and b and serum no. 4 in c and d.
Table III. ELISA inhibition with Lep d 2 variants. Concentrations of the inhibiting allergen needed to reach 50% inhibition of IgE binding to Lep d 2.0101 and Lep d 2.0102, and to Lep d 2.0201 and Lep d 2.0202, in L. destructor positive sera.

A

<table>
<thead>
<tr>
<th>Serum no.</th>
<th>Lep d 2.0101 (µg mL⁻¹)</th>
<th>Lep d 2.0102 (µg mL⁻¹)</th>
<th>Lep d 2.0101 (µg mL⁻¹)</th>
<th>Lep d 2.0102 (µg mL⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.029</td>
<td>0.006</td>
<td>0.020</td>
<td>0.007</td>
</tr>
<tr>
<td>2</td>
<td>0.100</td>
<td>0.045</td>
<td>0.081</td>
<td>0.048</td>
</tr>
<tr>
<td>3</td>
<td>0.195</td>
<td>0.090</td>
<td>0.206</td>
<td>0.215</td>
</tr>
<tr>
<td>4</td>
<td>0.085</td>
<td>0.074</td>
<td>0.083</td>
<td>0.074</td>
</tr>
<tr>
<td>5</td>
<td>0.238</td>
<td>0.062</td>
<td>0.206</td>
<td>0.064</td>
</tr>
<tr>
<td>6</td>
<td>0.471</td>
<td>0.091</td>
<td>0.424</td>
<td>0.095</td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th>Serum no.</th>
<th>Lep d 2.0201 (µg mL⁻¹)</th>
<th>Lep d 2.0202 (µg mL⁻¹)</th>
<th>Lep d 2.0201 (µg mL⁻¹)</th>
<th>Lep d 2.0202 (µg mL⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.014</td>
<td>0.004</td>
<td>0.008</td>
<td>0.004</td>
</tr>
<tr>
<td>2</td>
<td>0.055</td>
<td>0.043</td>
<td>0.051</td>
<td>0.043</td>
</tr>
<tr>
<td>3</td>
<td>0.058</td>
<td>0.047</td>
<td>0.055</td>
<td>0.047</td>
</tr>
<tr>
<td>4</td>
<td>0.079</td>
<td>0.064</td>
<td>0.085</td>
<td>0.070</td>
</tr>
<tr>
<td>5</td>
<td>0.047</td>
<td>0.029</td>
<td>0.047</td>
<td>0.030</td>
</tr>
<tr>
<td>6</td>
<td>0.060</td>
<td>0.049</td>
<td>0.062</td>
<td>0.051</td>
</tr>
</tbody>
</table>

To assess if the polymorphic residues have an effect on the type of T-cell response induced by Lep d 2, IFN-γ and IL-5 were measured in culture supernatants from long-term PBMC cultures. Although there were differences in the amounts of IFN-γ and IL-5 released by PBMC from individual patients, no consistent difference was found between the variants from either isoform in the six subjects investigated. However, the IFN-γ levels secreted after stimulation with the variants from the Lep d 2.02 isoform were slightly higher (median 16.3 ng·mL⁻¹; range 7.57-105) compared to after stimulation with the Lep d 2.01 variants (median 14.1 ng·mL⁻¹; range 3.08-80.0). The IL-5 secretion in response to the Lep d 2.01 variants (reaching a median level of 2.75 ng·mL⁻¹; range 0.42-6.05) was similar to that of the Lep d 2.02 variants (median 2.23 ng·mL⁻¹; range 0.53-3.93). A previous study has shown that Lep d 2 contains two immunodominant regions spanning amino acids 11-25 and 61-75 using peptides that represent all parts of the mature Lep d 2 (Eriksson et al. 2001). The fact that we did not see any difference in cytokine release among the Lep d 2 variants, that differing at amino acid 55 and 102, is therefore not surprising.

In conclusion, we have shown that both cultured and wild L. destructor mites contain the same pattern of polymorphisms. Furthermore, the Lep d 2 sequence diversity identified in this study does not seem to have any significant impact on the allergen’s IgE binding capacity or ability to induce cytokine release by T-cells.
5.2 DETERMINATION OF THE 3D STRUCTURE OF FEL D 1 (II, III)

We aimed to determine the 3D structure of the major cat allergen Fel d 1 to increase the understanding of the mechanisms behind cat allergy, and to obtain structural information for development of new forms of treatment.

In the initial crystallisation experiments of rFel d 1, protein that was purified using metal chelate affinity- and size exclusion chromatography was used. This resulted in an inadequate type of crystals, called spherulites. By using an additional purification step, ion exchange chromatography, an unknown contaminant could be removed and plate-shaped crystals were produced. Recombinant Fel d 1 crystals appeared in a MPD grid screen (10-65% MPD, pH 4-9) at 4°C. The size and quality of the crystals was further enhanced by reducing the protein concentration to 2.5 mg·mL⁻¹ and by seeding using a cat whisker. The final crystals were obtained in 13% MPD, 0.1 M sodium acetate pH 4.8 at 4°C, Fig. 8. The SeMet-Fel d 1 crystals were produced by seeding from wild-type crystals in 16% MPD, 0.1 M sodium acetate pH 4.8, using a protein concentration of 2 mg·mL⁻¹. The 3D structure of rFel d 1 was determined to a resolution of 1.85 Å, table IV.

Figure 8. Recombinant Fel d 1 crystal.
Photograph of a typical crystal of rFel d 1.
The longest dimension is ~ 0.2 mm.

The structure of rFel d 1 represents a unique fold among allergens. The fold of rFel d 1 consists of eight helices, H1-H4 and H5-H8, that correspond to chain 2 and chain 1, respectively, in native Fel d 1, Fig. 9. The N-terminal part comprising H1-H4 packs against the C-terminal part of the chain, H5-H8, forming a globular molecule of dimensions 40 x 30 x 40 Å. Three disulphide bonds linking the N-terminal with the C-terminal half of the monomer are found between residues C7-C162, C48-C136 and C73-C95. Despite a sequence identity of only 9%, the parts corresponding to chain 1 and 2 in natural Fel d 1 are similar in structure, with an overall root mean square deviation (rmsd) of 1.7 Å for 67 equivalent Cα atoms. Among the few conserved amino acids between the two chains are the three cysteines that participate in the formation of the disulphide bridges. The conservation of the 3D structure in chain 1 and chain 2 and of the critical cysteine residues suggests a common evolutionary origin.
Fel d 1 occurs naturally as a tetramer composed of two identical heterodimers (Duffort et al. 1991), but it was difficult to define the dimer of rFel d 1 (that corresponds to the native tetramer) from the crystal structure. The sizes of the contact areas between either adjacent monomers in the asymmetric unit or symmetry related molecules are very similar, ranging from 420 to 470 Å², typical of crystal contacts rather than oligomer interfaces. Therefore, assignment of the natural tetramer is not possible, probably due to the presence of MPD in the crystallisation solution. However, by using the Fel d 1 (1+2) construct we have recently obtained crystals in PEG3350 diffracting to 1.65 Å resolution. Data processing and initial model building indicates that a dimer has been formed.
**Figure 9.** Overall structure of rFel d 1.
The monomer is displayed from two different orientations, with a rotation of about 90° around the vertical axis. The helices corresponding to chains 2 and 1 are coloured in blue and gold, respectively. The dotted line indicates the disordered loop (residues 75 to 92). The three disulphide bridges that link chains 1 and 2 are displayed in green. An arrow indicates the unique glycosylation site at residue N33.

**Figure 10.** Comparison of rFel d 1 and uteroglobin.
*a,* Structure-based alignment of rFel d 1 chain 2 and 1, and of uteroglobin from rabbit (accession no. CAA25669) and man (accession no. AAH04481). Identical residues are indicated in bold. The secondary structure elements are indicated above the sequence. *b,* Stereo view of a ribbon diagram of the main chain of rFel d 1 (blue) superimposed on the oxidised form of rabbit uteroglobin (PDB entry 2UTG) (yellow).
Despite an overall sequence identity of only 20% with human CC10/secretoglobin, the structure of rFel d 1 reveals a striking similarity to the 3D structure of rabbit uteroglobin/secretoglobin (PDB entry 2UTG) (Callebaut et al. 2000, Klug et al. 2000), Fig. 10. The first secretoglobin identified was found in the uterus and was therefore named uteroglobin. However, similar proteins have later been found in other tissues and species, for example CC10 and CC16 from Clara cells in the human lung. Today it is known that secretoglobins are secreted by virtually all mucosal epithelial cells in mammals. Since its discovery it has been given numerous names that were based on which organ it was detected in or the type of ligand it was connected with. Recently a nomenclature committee have agreed on the family name secretoglobin for these proteins (Klug et al. 2000). Interestingly, the human secretoglobin gene has been mapped to a region of chromosome 11 that had previously been linked to atopy and bronchial hyperresponsiveness (Doull et al. 1996).

Sequence comparisons of the first chain of Fel d 1 with members of the secretoglobin family indicate that chain 1 is related to this protein family (Klug et al. 2000). However, the sequence homology between chain 2 and CC10 (13% identity based on 3D-structural alignment) is barely detectable and has not been noted until recently (Callebaut et al. 2000). It was therefore unclear how similar the structure of rFel d 1 would be to that of secretoglobins. The overall rmsd for 124 equivalent Cα atoms after superposition of rFel d 1 and the uteroglobin dimer (PDB entry 2UTG) is 1.7 Å, while the rmsd of chain 1 (involving helices H1-H4) and of chain 2 (helices H5-H8) of rFel d 1 to uteroglobin are 1.2 Å and 1.1 Å, respectively. Two of the disulphide bridges are conserved between uteroglobin and Fel d 1: C7-C162 and C73-C95. Although the third disulphide bridge at position C48-C136 is not present in uteroglobin (Bally and Delettre 1989), the structures are very similar in the vicinity of these amino acids. In addition, it has been reported that in many cases residue V44 in uteroglobin (corresponding to C48 in Fel d 1) is exchanged to a cysteine (Callebaut et al. 2000).

Interestingly, an internal cavity is formed at the packing interface between the two halves of the molecule, and residues lining the cavity originate from practically all of the helices in the monomer. Strong residual electron density, which cannot be attributed to side chains or ordered water molecules, indicates a bound ligand within the cavity. The nature of the ligand in the recombinant protein is unknown, but it has been modelled as an MPD molecule and six water molecules. The shape of the cavity differs from the corresponding cavity within uteroglobin (PDB entry 2UTG), Fig. 11. The size of the pocket is smaller (480 Å³) than in uteroglobin (750 Å³). Furthermore, the Fel d 1 cavity is asymmetric. One part of the cavity is lined by only hydrophobic residues, whereas polar and charged residues dominate on the opposite site. The hydrophilic end is created by two aspartic acids, D60 and D101, and two polar tyrosines, Y49 and Y113, Fig. 11a. In contrast, the cavity in homodimeric uteroglobin is symmetrical, mainly composed of hydrophobic residues, with a tyrosine from each chain (Y21 and Y21’) at opposite ends of the cavity, Fig. 11b. It has been shown that secretoglobins are a carriers of small ligands such as PCB derivatives, phosphatidyl inositol, phosphatidyl choline, retinol and progesterone (Umland et al. 1994, Härd et al. 1995, Mukherjee et al. 1999). The tyrosine residue Y21 is involved in the binding of ligands to uteroglobin (Härd et al. 1995) and the mutation of this position to an alanine or to a phenylalanine
results in a dramatic decrease of progesterone binding to uteroglobin (Peter et al. 1991). The particular distribution of residues in the cavity of rFel d 1 provides an amphipathic character to the cavity, which may be related to the physico-chemical properties of the endogenous ligand.

The allergen Fel d 1 and the cytokine-like secretoglobin protein family do not merely share a similarity in 3D structure, Fig. 10b: the expression of both proteins in epithelial cells (Charpin et al. 1991, Miele et al. 1994) is controlled and induced by steroids (Charpin et al. 1994, Zielonka et al. 1994, Mukherjee et al. 1999). The structural similarities of the two proteins suggest that anti-Fel d 1 antibodies might cross-react with uteroglobin. Hypothetically this cross-reaction might decrease the anti-inflammatory properties of the latter (Mukherjee et al. 1999), and thereby aggravating the allergic disease. Despite the low sequence identity between Fel d 1 and uteroglobin, three clusters composed of identical residues are present on the surface of the two molecules: cluster 1 (T109, P110 and D138) cluster 2 (E5, P8, K155 and L161) and cluster 3 (C73, C95, P96 and E93). These clusters may form the basis for cross-reactive epitopes in Fel d 1 and uteroglobin. The structural similarity between rFel d 1 and uteroglobin may also suggest that the allergen has cytokine-like properties, rendering it capable of modulating the immune response.

Previously three important IgE epitopes have been defined on Fel d 1 using 14-residue-long overlapping peptides spanning both chains. Two epitopes were found in chain 1 (residues 25-38, 46-59) and one in chain 2 (residues 15-28) (van Milligen et al. 1994). By using this information and the crystal structure of rFel d 1, solvent exposed residues available for IgE binding were identified, Q119, L123, P124, E128, A139, E143, E144 and E147 in chain 1 and residues F15, N19, E22, L23 and L27 in chain 2, Fig. 12.

In conclusion, rFel d 1 displays a striking structural similarity to uteroglobin, a steroid-inducible cytokine-like molecule with anti-inflammatory and immunomodulatory properties. The 3D structure of rFel d 1 provides a framework for further studies aimed at understanding the pathogenesis of cat allergy and creating new tools for diagnosis and therapy.
Figure 11. Comparison of the cavities in rFel d 1 and uteroglobin.

a. Representation of the cavity inside the rFel d 1 allergen and the residues lining the rFel d 1 cavity. b. Representation of the cavity within uteroglobin (PDB entry 2UTG) and the residues lining the uteroglobin cavity. The molecular surface of the cavities was calculated with a 1.4 Å probe radius and is coloured in blue. Hydrophobic, polar, acidic and basic residues lining the cavities within both molecules are coloured in grey, yellow, red and blue, respectively.

Figure 12. Predicted IgE epitopes.
The relative localisation of three IgE epitopes is indicated on the molecular surface of rFel d 1. The surface accessible residues of the epitopes 15-28 (chain 2) and 117-130, 138-151 (chain 1) are coloured in red, cyan and blue, respectively. a. Same orientation as in figure 10b. b. Perpendicular view to figure 12a. c. Perpendicular view to figure 12b.
5.3 RATIONAL DESIGN OF HYPOALLERGENS (IV)

In this study we present a novel approach, in which knowledge about an allergen structure and its B- and T-cell epitopes is used to rationally design hypoallergens. The idea is to introduce duplications of T-cell epitopes (DTE), preferably in a known B-cell epitope, in order to destabilise the 3D structure and simultaneously preserve the T-cell reactivity of the allergen. In addition, one can use directed mutagenesis of selected amino acid residues, an established method to reduce the IgE reactivity of allergens, to further reduce the IgE reactivity. To evaluate the feasibility of this combined approach, we have used the major cat allergen Fel d 1 as a model allergen.

We chose to duplicate the sequences encoding amino acid residues 23-30 on chain 1 and 46-53 on chain 2 in the rFel d 1 construct. These sequence segments have earlier been shown to induce frequent T-cell responses in a panel of T-cell lines (Counsell et al. 1996). The two sequences selected for DTE are parts of solvent-exposed loops between helix 2 and 3 of chain 2, and helix 6 and 7 of chain 1 (Kaiser et al. 2003a, Kaiser et al. 2003b) known to harbour IgE-binding epitopes (van Milligen et al. 1994). Interestingly, the role of loops as antibody binding epitopes has recently been demonstrated in several aeroallergens, cedar tree allergen Jun a 3 (Soman et al. 2000), Bet v 2 (Fedorov et al. 1997), Der p 2 (Hakkaart et al. 1998) and ragweed pollen allergen Amb t 5 (Rafnar et al. 1998), as well as in a detailed study of an allergen-antibody complex (Mirza et al. 2000). The Fel d 1 construct containing the duplicated (D) sequences was named rFel d 1 (D). In a second step, disulphide bonds in rFel d 1 and rFel d 1 (D) were disrupted by changing cysteine residues to serines. Mutagenesis was performed on cysteines at position 44 on chain 1, and at positions 7 and 48 on chain 2. In this way, molecules with one (Fel d 1 (I)), two (Fel d 1 (II) or three (Fel d 1 (III)) cysteines replaced by serines were created. In total, an array of seven Fel d 1 hypoallergen candidates was systematically engineered, table V and Fig. 13.

Table V. Fel d 1 derivatives.

<table>
<thead>
<tr>
<th>Name</th>
<th>Modification</th>
</tr>
</thead>
<tbody>
<tr>
<td>rFel d 1 (I)</td>
<td>C44S chain 1</td>
</tr>
<tr>
<td>rFel d 1 (II)</td>
<td>C44S chain 1, C48S chain 2</td>
</tr>
<tr>
<td>rFel d 1 (III)</td>
<td>C44S chain 1, C48S chain 2, C7S chain 2</td>
</tr>
<tr>
<td>rFel d 1 (D)</td>
<td>DTE</td>
</tr>
<tr>
<td>rFel d 1 (D I)</td>
<td>DTE, C44S chain 1</td>
</tr>
<tr>
<td>rFel d 1 (D II)</td>
<td>DTE, C44S chain 1, C48S chain 2</td>
</tr>
<tr>
<td>rFel d 1 (D III)</td>
<td>DTE, C44S chain 1, C48S chain 2, C7S chain 2</td>
</tr>
</tbody>
</table>
Altogether, three types of constructs were made and analysed for IgE binding in this study. The first type was a Fel d 1 derivative with DTE but without cysteine mutations, rFel d 1 (D). The IgE reactivity of rFel d 1 (D) was assayed in ELISA inhibition experiments and was shown to be considerably lower than that of rFel d 1. However, a more extensive investigation was not possible, since this construct was expressed in limited amounts and was mainly recovered in an aggregated form. A reason for the unsuccessful folding might be that a disulphide bond is part of the duplicated sequence of chain 2 and that this interferes with the folding process. The second type of construct was Fel d 1 with one or two disulphide bonds broken but without DTE, rFel d 1 (I), rFel d 1 (II) and rFel d 1 (III). These derivatives were easily expressed and purified as dimers. However, they displayed only moderate reduction in IgE-binding in the ELISA inhibition assay and were therefore not studied further. Hence, disruption of the disulphide bonds had only minor effects on the overall fold of the allergen. This is consistent with results from a comparison between the 3D structure of reduced and of oxidised uteroglobin (Härd et al. 1995). Reduction left the overall structure of uteroglobin intact, with only small changes seen at the terminal regions. The third type
of Fel d 1 construct, derivatives with both duplications and one or two disulphide bonds broken (rFel d 1 (D I), rFel d 1 (D II) and rFel d 1 (D III)), were easily soluble. In addition, they exhibited considerably reduced IgE-binding capacity (400 to 900 times) and were selected for further characterisation. The concentrations required for 50% inhibition of IgE binding to rFel d 1 in competition ELISA were 21.8 µg·mL⁻¹ for rFel d 1 (D I), 14.3 µg·mL⁻¹ for rFel d 1 (D II) and >30 µg·mL⁻¹ for rFel d 1 (D III). For rFel d 1, a concentration of 0.034 µg·mL⁻¹ was needed to reach 50% inhibition. The biological activity of the three Fel d 1 derivatives was compared to rFel d 1 by evaluating their ability to degranulate basophils. As shown in figure 14, a dose-dependent activation of basophils was seen in all four cat-allergic patients after stimulation with all derivatives as well as with rFel d 1. Generally, a lower concentration of rFel d 1 was needed to reach the same level of activation as achieved by the derivatives. The ability to induce degranulation of basophils decreased with increasing number of cysteine mutations in the derivatives. Accordingly, rFel d 1 (D III) showed the lowest capacity to activate basophils, Fig. 14.

The lymphoproliferative activity of rFel d 1, rFel d 1 (D I), rFel d 1 (D II) and rFel d 1 (D III) was assayed in PBMC from nine cat-allergic patients. PBMC from seven patients were stimulated to proliferate (SI > 2.0) in response to the different Fel d 1 proteins, whereas no proliferation was detected in PBMC from two of the cat-allergic patients. The presence of “non-responders” is in agreement with other studies regarding allergen stimulation of PBMC (Arquint et al. 1999, Eriksson et al. 2001). The SI values

---

**Figure 14.** Reduction in dose-dependent activation of basophils by Fel d 1 derivatives. Whole blood cells from four cat-allergic patients (A1-A4) were stimulated with serial dilutions of rFel d 1 and the derivatives rFel d 1 (D I), rFel d 1 (D II) and rFel d 1 (D III). Activated basophils are represented by the percentage of CD63⁺-cells among the gated CD203c⁺-cells.
measured in PBMC from the responding patients ranged 2.0-40.1 for rFel d 1, 2.8-162.9 for rFel d 1 (D I), 3.0-190.5 for rFel d 1 (D II), and 2.4-151.4 for rFel d 1 (D III). In fact, the hypoallergens had a stronger stimulatory effect than rFel d 1 in six of the seven responding patients, which could be due to the duplication of a T-cell epitope in each chain of Fel d 1. Only a weak response to rFel d 1 (D I) (SI 2.3), but not to the other Fel d 1 molecules, was observed in PBMC from a non-atopic control.

In conclusion, we have created three hypoallergens of Fel d 1 by using DTE in combination with directed mutagenesis of critical amino acid residues. When applied to the cat allergen Fel d 1, DTE accounted for reduction in IgE-binding capacity whereas mutagenesis improved protein refolding and solubility. The novel concept of DTE may be a valuable approach for the design of new hypoallergenic derivatives of common allergens.
6 FUTURE PERSPECTIVES

This thesis has focused on proteins/allergens that cause allergic sensitisation and disease. By use of recombinant allergens, clues to the mechanisms behind allergic sensitisation and suggestions for improved diagnostics and therapy have been provided. In addition, the studies have raised a number of new questions.

Our aim in study (III) was to determine the structure of recombinant Fel d 1. However, only the protein monomer was identified even though rFel d 1 was purified as a dimer, corresponding to the natural tetramer. We hope that by using another type of Fel d 1 construct for expression and, maybe more importantly, using a crystallisation condition that can preserve the quaternary structure of the protein, the dimer can be identified. By applying this approach we have recently obtained some promising results indicating that we have crystallised a dimer. The complete structure of Fel d 1 may provide us not only with the knowledge we need to design improved molecules for treatment, but also with basis for a detailed study of B-cell epitopes.

It has been shown that the human and rabbit secretoglobins bind hydrophobic molecules (Umland et al. 1994, Härd et al. 1995, Mukherjee et al. 1999). Consequently, several biological functions have been proposed for these secretoglobins related to their capability to harbour ligands (Umland and Sax 1995). However, we have shown that exactly the same molecules are not possible ligands of Fel d 1 for reasons discussed in this thesis. Consequently, human and rabbit secretoglobins and Fel d 1 must have somewhat different functions. In order to investigate the ligand and ligand binding of Fel d 1 we have started to isolate natural Fel d 1 from cat fur. Furthermore, it would be interesting to study the interaction between Fel d 1 and other proteins to learn more about why Fel d 1 is such a potent allergen. No receptors have been found for Fel d 1 or other secretoglobins, and by panning a human lung phage-display library with Fel d 1 and uteroglobin we hope to identify proteins in the lung that bind uteroglobin and Fel d 1.

The properties of the Fel d 1 derivatives generated in (IV) are typical for hypoallergens and are therefore promising new tools for treatment of cat-allergic patients. However, the efficacy of the hypoallergens has to be evaluated in different animal allergy models. In addition, before the derivatives can be used in SIT on humans, they should also be tested for safety in animals.

The availability of well characterised recombinant allergens has during the last decades provided us with the means to carefully study and genetically modify allergens. This has led to the first studies using recombinant birch hypoallergens in SIT (Gafvelin et al. submitted, Niederberger et al. submitted). Our results and those of others will hopefully lead to more efficient and safer diagnostics and therapy for allergic patients.
Allergi är ett växande problem i hela världen och i Skandinavien är proteiner från katt och kvalster de vanligaste orsakerna till allergi. För att kunna förbättra både diagnostik och behandling av allergisjukdomar krävs en noggrann karakterisering på molekylär nivå av de enskilda proteinerna, de så kallade allergenen, som allergiker reagerar mot. Avhandlingen är uppbyggd av fyra studier som behandlar karakterisering av två olika allergen, Lep d 2 som är det vanligaste allergenet från kvalstret *Lepidoglyphus destructor* (Studie I) och Fel d 1 som är det vanligaste och mest potenta allergenet från katt, *Felis domesticus* (Studier II, III). I avhandlingen ingår också ett arbete som tar upp en ny metod för att skapa förbättrad behandling av allergi, där Fel d 1 har använts som modell (Studie IV).

**Studie I**


**Studie II, III**

Målsättningen med den andra och tredje studien var att bestämma Fel d 1s tredimensionella struktur. Ett proteins tredimensionella struktur kan studeras med röntgenkristallografi. En kristall är uppbyggd av en enhet som upprepas väldigt många gånger, liksom en tegelvägg uppbyggd av tegelstenar, och en proteinkristall innehåller minst en proteinmolekyl i varje enhet. När man utsätter en kristall för röntgenstrålning kommer strålarna att spridas i ett bestämt mönster när de träffar elektronerna i kristallen och upprepningen av molekylen förstärker mönstret så att det kan mätas. Spridningen av röntgenstrålarna bearbetas sedan matematiskt och en tredimensionell modell av proteinet kan ritas upp. När vi bestämde strukturen visade sig Fel d 1 vara mycket likt ett annat protein, uteroglobin, som man tror har förmågan att modulera immunförsvaret. Inuti Fel d 1 fann vi också en ficka där en okänd molekyl var gömd. Vi spekulerar i att likheten med uteroglobin kan vara orsaken till att just proteinet Fel d 1, och inte något annat av de tusentals proteiner som katten producerar, är ett allergen. Genom att använda kunskaperna om strukturen av Fel d 1 samt utnyttja tidigare publicerad information om var på Fel d 1 IgE binder, kunde vi också bestämma vilka aminosyror som kan vara inkludade i bindningen till IgE. Kunskapen om strukturen av Fel d 1 har inte bara hjälpt oss att börja förstå varför proteinet är ett allergen, den är också en grund för att kunna förbättra behandling av kattallergiker.
**Studie IV**

Immuntärapei är den enda form av behandling av allergi som har en långvarig effekt. För att skapa tolerans mot ett allergen injiceras ökande doser av allergen under huden regelbundet i upp till fem års tid. Tyvärr kan olika typer av biverkningar förekomma. Målet med den fjärde studien var att pröva en ny metod för att skapa allergen med mindre IgE-bindande förmåga och därmed minska biverkningarna av immuntärapein. Vi valde att använda Fel d 1 som modellallergen eftersom vi har kunskaper om dess tredimensionella struktur, var på allergenets IgE binder och vilka strukturer som orsakar aktivering av immunförsvaret. Sju olika Fel d 1-derivat producerades och efter att ha testats för bl.a. minskad IgE-binding, valdes tre derivat ut som visade upp till 900 gånger lägre IgE-bindning än naturligt Fel d 1. Dessa proteiner är nu kandidatproteiner för användning i immuntärapei, men innan dess måste de utvärderas i olika djurmodeller.
8 ACKNOWLEDGEMENTS

Several people have contributed and supported me during my thesis project.

First of all, I would like to thank my two supervisors, Marianne van Hage-Hamsten and Guro Gafvelin, for accepting me as your PhD-student and introducing me to medical science and research. You two have been great, showing incredible patient, endurance and trust.

Omid Rasool, my co-supervisor during my first years in the lab, thank you for sharing your large knowledge in molecular biology and for a lot of fun.

I would also like to thank Neda Bigdeli for great technical assistance in the lab and Gerd Franzon Lundkvist for secretarial help.

And, a big thank you to all the other members in both Annika´s and Marianne´s groups, former and present, for creating a very nice atmosphere. I would like to acknowledge Tove, Hans, Tiiu and Sarah (in order of appearance) in particular. Thank you for all help, from spelling to chromatography, and great friendship, from pep talks to biking, badminton and sailing in the Greek archipelago.

Adnane Achour, Gunter Schneider and Tanja Sandalova, thank you for introducing me into crystallography and for letting me be part of your labs. I would also like to thank Stina, Anna, Jenny, Stefano and Lucas at MBB for all help and for inviting me to your parties!

Finally, I am thankful to my friends and my family for your encouragement and interest in what I “actually” am doing. Last but not least, Love, for your endless support and trust in me.

Tack/ Lotta
9 REFERENCES


Adedoyin J, Grönlund H & van Hage-Hamsten M. (submitted) Immunoglobulins, a new family of cat allergens including IgA and IgM, major allergens with mutually cross-reactive epitopes.


Bousquet J, Lockey R & Malling HJ. (1998) Allergen immunotherapy: therapeutic vaccines for allergic


Furmonaviciene R & Shakib F. (2001) The molecular basis of allergenicity:


cat allergen, cystatin (Fel d 3), a cysteine protease inhibitor. *Clin Exp Allergy* 31:1279-1286.


Molecular characterisation of major allergens from mite and cat


Development of a novel allergy vaccine: Genetically modified allergens prevent progression of allergic disease.


Liselotte Kaiser


Molecular characterisation of major allergens from mite and cat


