Analysis of the IgE epitope profile of the soybean allergen Gly m 4

Zur Erlangung des akademischen Grades eines Doctor rerum naturalium (Dr. rer. nat.) genehmigte Dissertation vorgelegt von Felix Husslik (M.Sc. Biochemie und Molekulare Biologie)

TECHNISCHE UNIVERSITÄT DARMSTADT

Analysis of the IgE epitope profile of the soybean allergen Gly m 4

Vom Fachbereich Chemie

der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades eines

Doctor rerum naturalium (Dr. rer. nat.)

genehmigte

Dissertation

vorgelegt von

Felix Husslik (M.Sc. Biochemie und Molekulare Biologie)

aus Miltenberg

Referent:	Prof. Dr. Harald Kolmar
Korreferent:	Prof. Dr. Stefan Vieths
Tag der Einreichung:	21. Juli 2016
Tag der mündlichen Prüfung:	25. Oktober 2016

Darmstadt 2016

D 17

Die vorliegende Arbeit wurde zwischen November 2011 und Dezember 2013 in der Forschungsgruppe „Molekulare Allergologie“ und zwischen Januar 2014 und November 2014 im Fachgebiet „Rekombinante Allergentherapeutika“ in der Abteilung Allergologie des Paul-Ehrlich-Instituts unter Betreuung von Dr. Dirk Schiller durchgeführt und am Fachbereich Chemie der TU Darmstadt als Dissertation angenommen.

Dekan: Prof. Dr. Gerd Buntkowsky

Referent: Prof. Dr. Harald Kolmar
Korreferent: Prof. Dr. Stefan Vieths

Datum der Einreichung: 21. Juli 2016

Datum der Disputation: 25. Oktober 2016

Table of contents

ABBREVIATIONS

°C	degree Celsius
µ	micro
A	Adenine
Å	Ångstrom
aa	amino acid(s)
Api g	*Apium graveolens* (celery)
APS	ammonium persulfate
Ara h	*Arachis hypogaea* (peanut)
BCA	bicinchoninic acid
Bet v	*Betula verrucosa*
bp	base pair
BSA	bovine serum albumin
C	Cytosine
CD	circular dichroism/cluster of differentiation
cf	*confer*
Cor a	*Corylus avellana* (European hazel)
C-terminus	carboxy terminus
CV	column volume
Da	Dalton
DNA	deoxyribonucleic acid
Dau c	*Daucus carota* (wild carrot)
DBPCFC	double-blind placebo-controlled food challenge
dH_2O	distilled H_2O
DLS	dynamic light scattering
dNTP	deoxynucleotide
DTT	dithiothreitol
E. coli	*Escherichia coli*
EC_{50}	half maximal effective concentration
EDTA	ethylendiamine tetraacetic acid
e. g.	*exempli gratia* (for example)
ELISA	enzyme-linked immuno sorbent assay
et al.	*et alia* (and others)
FPLC	fast performance liquid chromatography
G	Guanine
g	gram/gravitational constant (9.81 m/sec^2)
Gly m	*Glycine max* (soybean)
h	hour(s)
HEPES	4-(2-hydroxyethyl)-1-piperazine-ethansulfonic acid
His-tag	hexahistidine tag
HPSF	High purity salt free
HRP	Horse-radish peroxidase
Ig	immunoglobulin
IL	interleukin

IMAC	immobilized metal ion affinity chromatography
IPTG	isopropyl-β-D-thiogalactopyranoside
IUIS	International Union of Immunological Societies
kU	kilo units
LB	Luria Bertani
mAb	monoclonal antibody
Mal d	*Malus domestica* (apple)
min	minute(s)
MOPS	3-(N-morpholino)propanesulfonic acid
MS	mass spectrometry
MWCO	molecular weight cut-off
n.d.	not determined
NCS	(S)-Norcoclaurine Synthase
NMR	nuclear magnetic resonance
No.	number
NSB	non-specific binding
NTA	nitrilotriacetic acid
N-terminus	amino terminus
OAS	oral allergy syndrome
OD	optical density
p.a.	*pro analysi* (pro analysis)
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PEI	Paul-Ehrlich-Institut
Pen a	*Penaeus aztecus* (shrimp)
PMSF	phenylmethylsulfonyl fluoride
ppm	parts per million
Pru av	*Prunus avium* (cherry)
r	recombinant
RBL	rat basophilic leukaemia
rpm	revolutions per minute
RT	room temperature
SDS	sodium dodecylsulfate
SEC	size exclusion chromatography
SIT	specific immunotherapy
SOC	super optimal broth
T	Thymine
TAE	Tris-acetate-EDTA buffer
TBS	Tris buffered saline
TMB	3,3'-5,5'-tetramethylbenzidin
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet

One letter code	Three letter code	Amino acid
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartate/Aspartic acid
E	Glu	Glutamate/Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

1 INTRODUCTION

1.1 IgE-mediated allergy

The prevalence of pollen and food allergy is increasing not only in developed but also in developing countries worldwide (Campbell & Mehr, 2015). Allergic symptoms like rhinitis, urticaria or anaphylaxis affect up to 370 million individuals (estimated 50% of total population) in Europe leading to tremendous costs in health care systems (de Monchy *et al.*, 2013). Furthermore allergies have a great influence on patients' quality of life for instance through strict diets to avoid allergenic food or intake of medications during pollen season to reduce allergic symptoms. In Europe 150-300 million subjects (20-40%) are allergic to pollen (Vieths *et al.*, 2002; D'Amato *et al.*, 2007). Major elicitors are pollen from trees, grasses and weeds. Food allergy affects globally at least 5% of adults and 8% of children, however due to lack of detailed studies these numbers can be imprecise (Chafen *et al.*, 2010; Sicherer & Sampson, 2014). Recently it was shown that IgE sensitization to food allergens ranged from 7-24% of the European population depending on geographical location (Burney *et al.*, 2014). However mere presence of allergen-specific IgE antibodies does not necessarily correlate with allergic symptoms as subjects having allergen-specific IgE may well be clinically tolerant (Burney *et al.*, 2014). Overall eight food allergens cause the majority of food allergic reactions, particularly in children, which include allergens from plants such as peanuts, tree nuts, wheat and soy, together with allergens of animal origin including cow's milk, egg, fish and shellfish (Bush & Hefle, 1996; Allen & Koplin, 2012). In total, the self-reported prevalence of allergies against these foods is between 0.4% and 6.0% across all aged patients in Europe (Nwaru *et al.*, 2014). Prevalences can vary dependent on age, geographical location, dietary exposures or other factors (Sicherer, 2011).

In general an allergy is a hypersensitivity response of the immune system to normally innocuous proteins (Galli *et al.*, 2008). Food allergy but also pollen allergy can follow the mechanism of a type I allergy resulting in an immediate, IgE driven immune response (Rajan, 2003). Development of type I allergies can be divided into two phases (Figure 1).

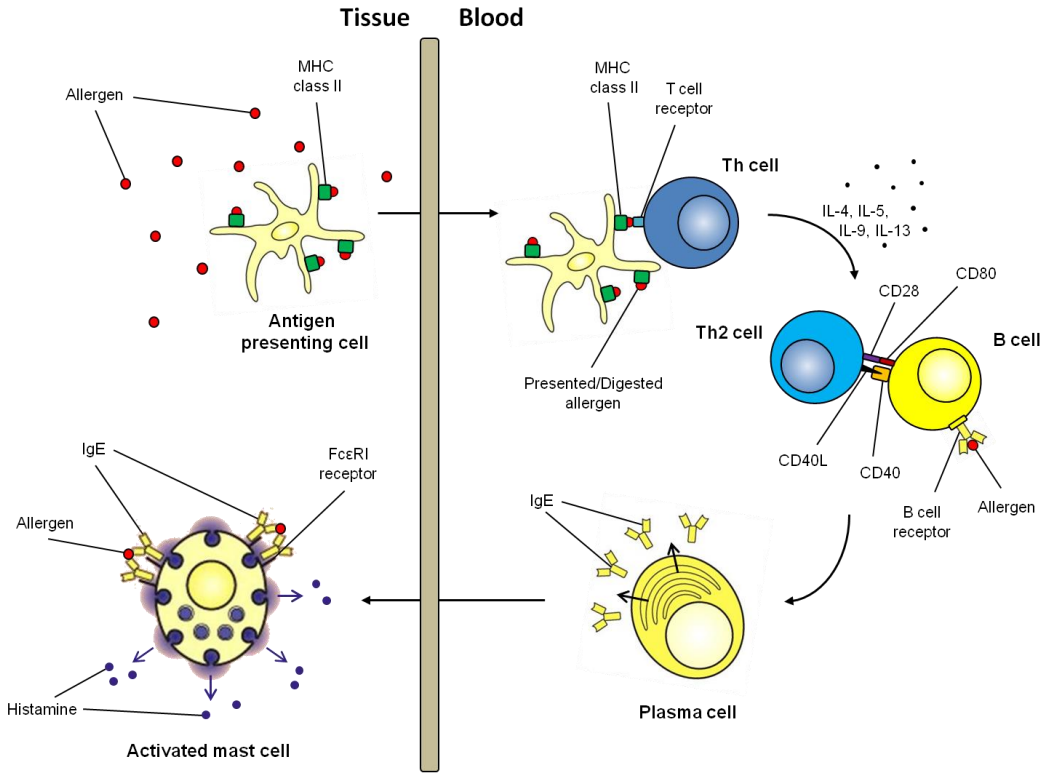


Figure 1: Pathomechanism of type I allergies.

For details see text. Adapted from Murphy *et al.*, 2012.

In the first phase, so called sensitization, an individual is exposed to an allergen without causing any allergic symptoms. Allergens can enter the immune system through either gastric or oral mucosa, by ingestion (food allergy) or inhalation (pollen allergy), respectively. Allergens are taken up and processed by antigen presenting cells (APC). APCs are dendritic cells, macrophages or B cells which present peptides of digested allergens on their surface bound to major histocompatibility complex (MHC) class II molecules. Naïve $CD4^+$ T helper (Th) cells circulating through blood stream and lymphoid organs can bind to peptide-MHC class II complexes via their appropriate T cell receptor resulting in proliferation of T cells and cytokine release (Murphy *et al.*, 2012). Besides interaction of T cell receptor with MHC class II molecules, additional allergen-unspecific signals including co-stimulatory molecules or cell-to-cell contacts are discussed for T cell differentiation (Beier *et al.*, 2007). Depending on resulting cytokine milieu a Th1 or Th2 driven immune response is established. Secretion of interferon (IFN)-gamma, interleukin (IL)-12 and tumor necrosis factor (TNF)-β is characteristic for Th1 cells while Th2 cells synthesize IL-4, IL-5, IL-9 and IL-13. In parallel B cells encounter their specific non-digested allergen and present it on their cell surface bound to their specific B cell receptor (Murphy *et*

al., 2012). These primed B cells can interact with activated Th2 cells via CD40-CD40 ligand and CD80-CD28 interactions, respectively, causing an isotype switch from B cells to plasma cells producing allergen specific IgE antibodies (Poulsen & Hummelshoj, 2007). In contrast to Th2 cells a Th1 driven immune response causes synthesis of IgG antibodies. Until now it is not known how the Th1/Th2 ratio causes different immune responses in atopic and non-atopic individuals (Romagnani, 2004). Secreted IgE antibodies can enter the blood stream and finally bind to the high affinity receptor for IgE (FCεRI) on the surface of effector cells like mast cells or basophil granulocytes (Galli & Tsai, 2012). If such allergen-specific IgE antibodies are present an individual is sensitized to the allergen.

In the second phase, the elicitation, a recurring contact with the allergen leads to allergen-induced IgE receptor complex formation. If the allergen presents at least two different antibody-binding sites, cross-linking of two membrane-bound IgE antibodies via allergen occurs. IgE cross linkage activates several signaling cascades and biochemical reactions resulting in degranulation and release of mediator molecules like histamine. Histamine immediately causes contraction of smooth muscle tissues, lowering of blood pressure or an increased permeability of blood vessels (White, 1990; Broide, 2001). This may lead to local inflammation like swelling and irritation of oral mucosa, gastrointestinal discomfort and sometimes also systemic reactions like anaphylaxis.

To date it is not known why some patients develop an allergy against normally harmless proteins while others do not. Several hypotheses what might contribute to development of allergies are discussed but underlying mechanisms are unknown. Often mentioned in this context is the “hygiene hypothesis” indicating a reduction of allergic reactions caused by exposure to infections during childhood (Strachan, 1989; Kay, 2001). Furthermore westernized lifestyle with change in medical treatment and diet, but also genetic factors, air pollution or stress as elicitors for allergies are discussed (Wills-Karp *et al.*, 2001; Saxon & Diaz-Sanchez, 2005; Vercelli, 2008).

1.2 IgE-mediated food allergy

IgE-mediated food allergy can be divided in class I and class II food allergic reactions. Class I food allergy, classical food allergy, mainly affects young children whereas class II food allergy is present in adults (Egger *et al.*, 2006). Allergic

reactions occur via sensitization through gastrointestinal tract in class I allergy or because of cross reactions between several food and pollen allergens in class II allergy, respectively (Breiteneder & Ebner, 2000). The latter is known as pollen-related food allergy. Molecular reason for pollen-food allergy can be cross reactivity of IgE antibodies against pollen allergens with structural homologous proteins from plant food (Kazemi-Shirazi *et al.*, 2000). Examples are cross reactions between birch pollen and several foods, celery-mugwort-spice and grass or ragweed pollen with fruits (Aalberse *et al.*, 2001). Patients with pollen-related food allergies are sensitized to pollen allergen and can exhibit different allergic symptoms after consumption of plant-derived foods and food products. Frequently only weak local reactions like itching and swelling of tongue or lips occur, but also severe, systemic symptoms such as anaphylaxis and urticaria are observed usually within 2 hours after food intake (Ballmer-Weber *et al.*, 2000; Ballmer-Weber *et al.*, 2001; Kleine-Tebbe *et al.*, 2002; Mittag *et al.*, 2004a; Mari *et al.*, 2005). As shown for type I allergies, inhalation of pollen allergens causes sensitization, including production of specific IgE antibodies. Afterwards the first oral contact with the food can be sufficient to cause an allergic reaction without prior sensitization to the specific food allergen. In birch pollen-related food allergy cross reactivity occurs between IgE directed against major birch pollen allergen (Bet v 1) and structurally homologous food allergens from various fruits like apple, apricot, cherry, peach, pear or strawberry as well as different nuts and legumes like peanut, hazelnut, mungbean or soybean, respectively, but also vegetables like carrots, celery or tomatoes can cause allergic symptoms (Vieths *et al.*, 2002; Breiteneder & Mills, 2006). Both, Bet v 1 and Bet v 1-related food allergens, show similar IgE binding regions (epitopes) on their molecular surface causing antibody cross reactions. In Europe up to 70 million subjects exhibit a birch pollen allergy and up to 90% of patients with birch pollinosis also show a birch pollen-related food allergy (D'Amato *et al.*, 2007; Geroldinger-Simic *et al.*, 2011). Bet v 1 and its homologous food allergens belong to the pathogenesis related protein 10 (PR-10) family. PR-10 proteins are abundant in higher plants. Expression of PR-10 proteins in the cytosol is induced in plants upon stress conditions like wounding, microbial or fungal elicitors and several abiotic factors (Somssich *et al.*, 1986; Warner *et al.*, 1992; Walter *et al.*, 1996; Robert *et al.*, 2001). Furthermore they play a role in plant development (Hoffmann-Sommergruber, 2000). For Bet v 1 ribonuclease activity and, together with its ligand quercetin-3-*O*-sophoroside, a protective function against UV

damage was suggested (Bufe *et al.*, 1996; Seutter von Loetzen *et al.*, 2014). Furthermore for Pru av 1 from cherry and Bet v 1 binding of phytosteroids was suggested but specific functions for other PR-10 proteins remain elusive (Neudecker *et al.*, 2001; Radauer *et al.*, 2008). Besides a high amino acid sequence identity (38-67%) with Bet v 1, homologous allergens show also an identical protein folding (Vieths *et al.*, 2002). Furthermore all allergens of the Bet v 1-family have a conserved glycine-rich P-loop region comprising amino acids 46-52 in Bet v 1 characterized as structural element of nucleotide binding proteins and important for IgE reactivity (Saraste *et al.*, 1990; Mirza *et al.*, 2000; Holm *et al.*, 2004). Related allergens are for example Mal d 1 (apple), Pru av 1 (cherry), Cor a 1.04 (hazelnut), Ara h 8 (peanut), Gly m 4 (soybean), Dau c 1 (carrot) or Api g 1 (celery). Besides IgE cross reactivity, these food allergens can also stimulate Bet v 1-specific T cells (Fritsch *et al.*, 1998; Jahn-Schmid *et al.*, 2005). In contrast to B cell activation, T cell cross reactivity even persists after gastrointestinal digestion (Schimek *et al.*, 2005). Furthermore, complexity of different food matrices can have a huge effect on stability and presentation of allergens (Grimshaw *et al.*, 2003).

1.3 Diagnosis and treatment of allergies

For correct treatment of allergic symptoms a precise and distinct diagnosis is essential. For this purpose different diagnostic procedures are used today. First conclusions on the potential presence of allergies can be drawn by a thorough anamnesis usually in combination with a standardized questionnaire and patient's history. Then the anamnesis needs to be verified by allergy tests. Very common are skin prick test (SPT) and prick to prick test (PPT). In both tests small amounts of allergen are applied to the skin and the size of resulting wheals is measured. Besides, *in vitro* tests are used to determine allergen-specific IgE antibodies in patients' sera. One example is ImmunoCAP™ blood test (Thermo Fisher, Sweden) a standard diagnostic tool used in the majority of clinical laboratories. ImmunoCAP™ results show levels of total IgE antibodies or levels of IgE specific for allergen extracts or individual allergen molecules. Positive results in both, *in vitro* analyses and *in vivo* skin tests only point to sensitization to pollen or food allergens but not necessarily to an allergy. Especially for diagnosis of food allergies double-blind placebo-controlled food challenge (DBPCFC), first described in 1976, is considered

as gold standard (May, 1976). In DBPCFC the suspicious food allergen or allergenic food is administered in increasing doses and allergic symptoms of the patients are recorded. For this, clinical monitoring and safety precaution are required making DBPCFC very expensive and time-consuming. In general only hospitals are able to perform such types of food challenges. In contrast to skin testing and *in vitro* testing, not only sensitization but confirmation of the clinical relevance of food allergies is possible. Furthermore DBPCFC can be used to judge the severity of a food allergy while to date no *in vitro* test is available to predict severity of allergic symptoms in patients (Spergel *et al.*, 2004).

Currently allergen avoidance is the only causal treatment of food allergies. However food avoidance can only be guaranteed by correct food labeling and exclusion of hidden allergens. Treatment of allergic symptoms is possible in both pollen and food allergies. Medications like antihistamines, epinephrine, corticosteroids, H2 blockers or β2-antagonists can reduce allergic symptoms immediately (Sicherer & Teuber, 2004). For about 100 years specific immunotherapy (SIT) is used as treatment of allergic symptoms caused by weed, grass or tree pollen, as well as house dust mite, cat dander or insect venom (Durham & Nelson, 2011). During SIT increasing dosages of sensitizing allergen are administered repeatedly in the so called updosing phase followed by maintenance phase with same dose injected monthly over 3 to 5 years. Nowadays subcutaneous injections, as well as sublingual administration are possible. Goal of SIT is to shift allergic Th2 response to a more tolerant Th1-driven immune response and therefore decrease of IgE-mediated allergic reactions. Nevertheless the overall mechanism behind SIT is not completely understood. During SIT specific IgE levels increase first, but decrease after several months of treatment accompanied with the switch of B cells to produce IgG antibodies mainly of the subtype IgG_4 (Aalberse *et al.*, 1983; van Ree *et al.*, 1997; Gehlhar *et al.*, 1999). These IgG antibodies act as so called blocking antibodies inhibiting allergen-IgE interactions by competition with IgE, resulting in a reduced allergic response (James *et al.*, 2011; Jutel *et al.*, 2013). Furthermore an increase in regulatory T cells is observable which produce anti-inflammatory cytokines like IL-10 and TGF-β and suppress formation of dendritic cells (Jutel *et al.*, 2003). IL-10 effects downregulation of Th2 cells and induces IgG_4 and IgA antibodies (Akdis *et al.*, 1998). TGF-β inhibits proliferation and differentiation of B cells and antibody response.

SIT for treatment of food allergies was tested as sublingual immunotherapy (SLIT)

where the food allergen in solution is held under the tongue for some time and oral immunotherapy (OIT) where increasing doses of food allergen is administered orally (Syed *et al.*, 2013). Both SLIT and OIT are currently not routinely used for treatment of food allergy because of limited success rates, lack of sustained effects and severe side effects.

Major drawbacks of current SIT are besides cost-intensity and long-lasting treatment for the patient, the risk of systemic allergic reactions especially in OIT when treating food allergies with egg, peanut or milk (Jones *et al.*, 2009; Burks *et al.*, 2012; Vázquez-Ortiz *et al.*, 2013). In addition, parallel medication has to be monitored for each patient and use of SIT in patients with autoimmune disorders or malignancy might cause life-threatening side effects. Despite the described beneficial long-term effect of SIT with grass pollen for example, some patients however do not respond to treatment (Durham *et al.*, 2010). To date several new approaches are discussed which might be used in addition to SIT in the future (Akdis, 2014). For example recombinant allergens purified from bacterial expression systems are already used in clinical trials compared to crude allergen extracts, even though an authorization has not been granted to date (Cromwell *et al.*, 2011). Furthermore modifications in recombinant allergens can reduce side effects and at the same time ameliorate effectiveness of SIT. One example is the use of hypoallergenic variants, peptides, chimeric or fused proteins which have lost their conformational B cell epitopes but still show linear T cell binding sites (Pree *et al.*, 2007; Kahlert *et al.*, 2008; Moldaver & Larché, 2011; Schülke *et al.*, 2014). These variants show lower allergenicity but preserved T cell immunogenicity (Karamloo *et al.*, 2005). In addition, recombinant allergens modified with modular antigen transporter or hypoallergenic fusion proteins which induce IgG blocking antibodies are discussed to date (Martínez-Gómez *et al.*, 2009; Focke-Tejkl *et al.*, 2015). Furthermore several additional routes for treatment of allergies like epicutaneous and intralymphatic administration and use of specific adjuvants are discussed as potential new approaches for treatment of food allergies (Senti *et al.*, 2012; Hylander *et al.*, 2013).

The efficacy of pollen SIT on pollen-related food allergies was analyzed in several studies but outcomes are contradictory (Bolhaar *et al.*, 2004; Bucher *et al.*, 2004; Mauro *et al.*, 2011). Patients with SIT for treatment of birch pollinosis may show of reduction in oral allergy syndrome to apple or hazelnut but however often only a temporary improvement of the food allergy is observed.

1.4 Soybean allergy

During recent years an increase in consumption of soybean food products like soy sauce, tofu or soy milk in Europe was observed. Soybean (*Glycine max*) is a member of the *Fabaceae* family of plants and is classified to the group of legumes. It has a very high total protein content of 40%. Even though more than 20 IgE-binding proteins from soybean are known, only eight soybean proteins are accepted so far as allergens by the International Union of Immunological Societies (IUIS) Allergen Nomenclature Sub-Committee. Soybean hull proteins Gly m 1 and Gly m 2 were described as responsible for asthmatic reactions in Spain (González *et al.*, 1991; Codina *et al.*, 1997). Both function as aeroallergens causing respiratory soy allergy and are not present in soy seeds. Gly m 3 is a profilin, while Gly m 4 belongs to the PR-10 protein family (Crowell *et al.*, 1992; Rihs *et al.*, 1999). β-conglycinin Gly m 5 and glycinin Gly m 6 cause severe allergic reactions after ingestion and are suggested as markers for severe soybean allergy (Holzhauser *et al.*, 2009; Ito *et al.*, 2011). Recently two additional soy allergens Gly m 7 and Gly m 8 were identified (Ebisawa *et al.*, 2013).

Gly m 4, formerly named starvation-associated message 22 (SAM22), is a Bet v 1-homologous protein in soy causing birch pollen-related soy allergy and affects roughly 10% of patients with birch pollinosis (Mittag *et al.*, 2004b). As shown for several PR-10 food allergens Gly m 4 can cause mild but also severe allergic symptoms (Kleine-Tebbe *et al.*, 2002; Kosma *et al.*, 2011; De Swert *et al.*, 2012; El-Hifnawi *et al.*, 2012). Depending on food product and industrial processing Gly m 4 level ranges from 0-70 milligrams per kilogram while strong heating and roasting of soy products completely abolish IgE binding of Gly m 4 (Franck *et al.*, 2002; Mittag *et al.*, 2004b). Gly m 4 three-dimensional structure was solved by nuclear magnetic resonance (NMR) spectroscopy in 2003 (PDB code 2K7H) (Neudecker *et al.*, 2003b). Structural characteristics are a seven-stranded antiparallel β-sheet enwinding a long C-terminal α-helix and two short α-helices. A hydrophobic cleft formed inside the molecule shows possible ligand binding ability (Radauer *et al.*, 2008). Gly m 4 consists of 157 amino acids and has a molecular weight of 16.6 kDa. It shares an amino acid sequence identity of 48% with Bet v 1a. The biological function, as well as the IgE-binding regions of Gly m 4 are not known.

1.5 Epitopes

The molecular structure within an allergen molecule involved in antibody binding is called epitope and the corresponding surface of the antibody is designated as paratope. Epitopes are thought to be mainly conformational indicating the importance of structural homology between allergens in pollen-food allergy (Ferreira *et al.*, 1998; El-Manzalawy & Honavar, 2010). In conformational epitopes amino acids that are distributed throughout the primary structure of the allergen are localized on a particular region on the molecular surface of the allergen upon folding into the native protein conformation. In contrast sequential epitopes are comprised of a linear stretch of adjacent amino acids displaying an antibody-binding site (Van Regenmortel, 1989). T cell epitopes are always linear epitopes, as during an allergic reaction T cell receptors only bind to peptides presented by antigen-presenting cells. In contrast most B cell epitopes are conformational, therefore a well-folded allergen is required for epitope recognition via B cells or IgE antibodies (Pomés, 2010). Usage of peptide-based approaches e. g. phage-display technology often results in identification of allergen mimotopes. These are peptides mimicking an antibody-binding site of the allergen (Davies *et al.*, 2000). Furthermore IgE epitopes can be either structural or functional. Whereas a structural IgE epitope of an antigen is described as all amino acids of the interface that are in close proximity to the paratope of an IgE antibody, a functional IgE epitope only represents individual residues accounting for most of the binding energy to IgE (Dall'Acqua *et al.*, 1998). Therefore in a structural IgE epitope such amino acids which dominate the energetics of allergen-antibody binding are defined as a functional IgE epitope and substitution of these residues causes a decrease in binding affinity.

1.6 Epitope analysis

To date little is known about clinically relevant antibody-binding sites of allergens. Until know only three IgE-binding B cell epitopes and eleven B cell epitopes of allergens in total are verified (Pomés, 2010). With structural methods like X-ray crystallography and NMR spectroscopy IgE epitopes of β-lactoglobulin, Phl p 2 and Art v 1 as well as binding sites of Bet v 1, Api m 2, Bla g 2, Der f 1 and Der p 1, lysozyme, Der p 2 and Blo t 5 with Fab fragments of monoclonal IgG_1 antibodies could be identified, respectively (Amit *et al.*, 1986; Mirza *et al.*, 2000; Mueller *et al.*,

2001; Niemi *et al.*, 2007; Padavattan *et al.*, 2007; Li *et al.*, 2008; Naik *et al.*, 2008; Padavattan *et al.*, 2009; Razzera *et al.*, 2010; Chruszcz *et al.*, 2012). The only known structurally solved epitope of Bet v 1 is the binding site with the monoclonal murine IgG_1 Fab fragment BV16 (Mirza *et al.*, 2000). This discontinuous epitope involves 16 residues of Bet v 1 comprising a molecular surface area of Bet v 1 of 931 $Å^2$. The majority of the epitope encompasses amino acids 42-52 including the Bet v 1 P-loop region. It was shown that interaction between BV16 and Bet v 1 can reduce serum IgE binding, suggesting a competition of Bet v 1-specific serum IgE and monoclonal antibody for an overlapping epitope of Bet v 1 (Spangfort *et al.*, 1999; Spangfort *et al.*, 2003). For other allergens from Bet v 1-family including Gly m 4, no epitopes have been identified so far. However information on the potential involvement of single amino acids in antibody binding of several allergens is available. For the identification of such residues different approaches were used in the last years.

Generating mutational variants of allergens can be used to identify amino acids critical for Ig binding. For this, knowledge of protein structure is essential to choose appropriate candidate residues on the allergen surface for site-directed mutagenesis. These include mainly charged amino acids putatively crucial for allergen-antibody interaction. If variants show reduced antibody binding while the overall protein conformation is comparable to wild type allergen, the substituted amino acid might represent a functional epitope (see 1.5). Besides the time-consuming expression and purification of a large quantity of protein variants this method has another disadvantage because protein variants might lose their overall folding when structurally important residues are substituted. Therefore it is always necessary to verify native-like protein structures of generated allergen variants because a reduced Ig binding can be caused by two reasons: First, an amino acid essential for the native protein fold is replaced, and the modified three-dimensional structure of the allergen has lost a conformational epitope. Second, a residue not important for overall protein folding but critical for allergen-antibody interaction is mutated causing a reduction in Ig binding while protein structure is not affected. Mutational epitope analysis was used with several allergens including Api g 1, Bet v 1, Mal d 1 or Pru av 1 (Ferreira *et al.*, 1996; Scheurer *et al.*, 1999; Neudecker *et al.*, 2003a; Holm *et al.*, 2004; Ma *et al.*, 2006; Wangorsch *et al.*, 2007; Klinglmayr *et al.*, 2009).

The method of peptide microarrays to identify epitopes of allergens uses chemically synthesized overlapping peptides covering the complete amino acid sequence of the

allergen (Reese *et al.*, 2001). These peptides can be analyzed for antibody binding either in solution for example in enzyme-linked immuno sorbent assay (ELISA) measurements or bound to a solid surface in microarray format. Using peptides of 5-15 amino acids in length linear epitopes of allergens can be identified (Vereda *et al.*, 2010). However use of overlapping synthetic peptides needs to be further evaluated as maximum peptide length or different shifting offsets resulted in different identified linear epitopes (Steckelbroeck *et al.*, 2008).

Furthermore a combination of both experimental and computational methods for epitope mapping is possible. This includes usage of phage-display technology where a library of short peptides presented by bacterio-phages is used to identify antibody-binding peptides, so called mimotopes (see 1.5). However it is believed that mimotopes represent only in part an antibody-binding site. Mimotopes can be mapped on allergen surface by computer-based methods to identify putative epitopes of an allergen (Davies *et al.*, 2000). Phage-displayed IgE-binding peptides could be identified for several allergens including four PR-10 proteins (Mittag *et al.*, 2006). In Ara h 8, Bet v 1, Gly m 4 and Pru av 1 three IgE-binding regions with conserved amino acids were mapped.

In addition monoclonal antibodies (mAb) can be used for epitope mapping as well. Binding sites of allergen-specific mAb which interact with allergen-derived peptides were identified (Lebecque *et al.*, 1997; Gieras *et al.*, 2011). Furthermore inhibition experiments can indicate potential overlapping binding sites between mAb and serum IgE.

While the methods for epitope identification described here require high experimental effort and a lot of time, *in silico* epitope predictions might be used to overcome these disadvantages. Up to now several algorithms exist for bioinformatical prediction of IgE binding sites (El-Manzalawy & Honavar, 2010; Dall'Antonia *et al.*, 2014). Algorithms are based on the sequence or structure of known allergens and consider the physicochemical properties of amino acids, surface characteristics, distance values or solvent accessibility to predict epitopes. Examples are web services like ElliPro and DiscoTope or software programs like Mapitope (Haste Andersen *et al.*, 2006; Bublil *et al.*, 2007; Ponomarenko *et al.*, 2008). Furthermore web-based servers like EpiSearch or MimoPro are used to map mimotopes derived from phage-display experiments (Negi & Braun, 2009; Chen *et al.*, 2011). Because of the lack of information on IgE epitopes, *in silico* analysis still needs experimental verification to

confirm the suitability of computational tools for the prediction of antibody-binding sites.

X-ray and NMR studies of allergen-antibody complexes are the most straight-forward approach to identify epitopes. Both approaches provide high-resolution structures of complexes between allergen and antibody including molecular details of interaction. Furthermore both linear and conformational epitopes can be identified.

The identification of epitopes by reduced allergen-antibody binding of recombinant allergen variants has two potential drawbacks. First, polyclonality of serum IgE makes it difficult to differentiate individual IgE epitopes in the context of total serum IgE antibody repertoire. Thus measuring the lack of just one particular epitope in the background of overall IgE binding to the allergen is technically challenging. Second, allergen variants might have a non-native-like protein conformation which makes them improper for studying conformational epitopes. To overcome these disadvantages epitope grafting can be used. Here, functional IgE epitopes of an allergen are transferred to a non-IgE-binding protein to induce antibody interaction. Provided that protein folding remains unaffected, IgE-binding would directly point to these grafted residues. Epitope grafting is contradictory to mutational analysis of allergens where amino acid substitutions might lead to a reduced IgE binding rather than induced allergen-antibody interaction. Until now epitope grafting was used to create chimerics between two allergens of the same protein family or by transfer of sequential IgE epitopes of shrimp allergen Pen a 1 onto non-allergenic mouse tropomyosin (Albrecht *et al.*, 2009; Holm *et al.*, 2011; Gepp *et al.*, 2014). An alternative approach could be the use of a non-IgE-binding protein with allergen-like protein structure to study the impact of individual amino acids on IgE binding.

1.7 NCS system

Epitope grafting for identification of IgE-binding regions needs a structurally homologous protein with low IgE binding. One example in case of the Bet v 1-family is (S)-Norcoclaurin-Synthase (NCS) from meadow rue *Thalictrum flavum* (Berkner *et al.*, 2007). Showing Bet v 1-type protein folding as determined by X-ray crystallography, NCS belongs to PR-10 family of proteins (Ilari *et al.*, 2009). In contrast to most members of this protein family the biological function of NCS is known. It catalyzes the first step in biosynthesis of benzylisoquinoline alkaloid which

is condensation of dopamine and 4-hydroxyphenylacetaldehyde to (S)-norcoclaurine. (S)-norcoclaurine is the central precursor of all benzylisoquinoline alkaloids, a large group of secondary metabolites in plants (Samanani *et al.*, 2004). Besides its structural homology to proteins of the Bet v 1 family, NCS only shows a sequence identity between 10-30% with members of PR-10 protein family. The protein has a molecular weight of 21 kDa and 210 amino acids in total. It was first characterized extensively as Δ29NCS variant with a deletion of the N-terminal signal peptide by Berkner *et al.* in 2008 (Berkner *et al.*, 2008). This truncated variant still has additional amino acids on both N- and C-terminus of the protein compared to Bet v 1 and its homologous allergens. Interestingly, despite the structural homology with PR-10 proteins, recombinant Δ29NCS variant exhibits no to very weak binding of Bet v 1-specific IgE antibodies (Dissertation Berkner). Therefore Δ29NCS resembles a promising candidate for epitope grafting. With this approach putative epitopes from Bet v 1 or homologous allergens can be transferred to Δ29NCS to generate Δ29NCS variants that display individual allergen-derived epitopes in an allergen-type protein conformation. These protein variants can then be analyzed for serum IgE binding in order to obtain information of relevant epitopes of the respective allergen. Within Hanna Berkner's PhD thesis several residues involved in binding between Bet v 1 and BV16 epitope were grafted onto Δ29NCS. Resulting variants were analyzed according IgE binding with sera of birch pollen-allergic subjects. Two of these Δ29NCS variants showed induced IgE binding in some patients while still showing Bet v 1-related protein structure (Berkner *et al.*, 2014). Therefore NCS might resemble a promising and emerging tool for epitope grafting with high potential for identification of antibody-binding sites of Bet v 1 and related allergens.

1.8 Aim of the study

Data on clinically relevant antibody-binding sites of Bet v 1-homologous allergens is very limited but the detailed knowledge on epitopes might lead to epitope-resolved diagnosis and maybe therapy of allergy. With component-resolved diagnosis (CRD) a differentiation between allergy and tolerance is not possible but epitope-resolved *in vitro* diagnosis might bear the potential to identify correlations with the clinical phenotype. Central to answering this question is the identification of the IgE epitope

profile and a potential correlation between clinical phenotype and IgE sensitization pattern.

The present study aimed at determining the IgE epitope profile of Gly m 4 in birch pollen-allergic subjects with and without clinically relevant birch-associated allergy to soybean. For this, amino acids of Gly m 4 potentially involved in binding of IgE were identified by i) bioinformatic mapping of IgE-binding phage-displayed peptides and Gly m 4-specific peptides onto the molecular surface of Gly m 4, ii) generation of a protein-based epitope library utilizing recombinant variants of the Gly m 4-type model protein norcoclaurine synthase (NCS), and iii) generation of substitutional variants of recombinant Gly m 4 with reduced IgE binding capacities. The IgE epitope profiles obtained were supposed to be correlated to clinical symptoms and allergen-specific IgE levels of the study population.

Furthermore it was intended to analyze misfolded recombinant variants of both Bet v 1 and Gly m 4 via immunological and physicochemical assays to evaluate IgE binding capacities and quality of recombinant allergen preparations.

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 Hard- and software

For generation of diagrams Microsoft® Excel 2007 (Microsoft Corporation) and data analysis software GraFit 5.0.13 (Erithacus Software Ltd., Surrey, UK) was used. DNA sequences were analyzed with Serial Cloner 2.6.1 (SerialBasics, USA). Protein structures were visualized and analyzed with PyMol v0.99 (DeLano Scientific LLC, San Francisco, CA, USA) and Swiss-PdbViewer 4.0.1 (Guex & Peitsch, 1997) while PyMol was also used to create figures of proteins and molecules. Relevant pdb files of proteins were downloaded from RCSB Protein Data Bank (http://www.rcsb.org/pdb/) (Berman, 2000). Modeling of protein structure was carried out using Phyre2 protein structure prediction server (Kelley & Sternberg, 2009) with the respective primary sequences of the modeled proteins downloaded from UniProt database (www.uniprot.org). For alignment of amino acid sequences EMBOSS Needle Pairwise Alignment tool provided by European Bioinformatics Institute was used (http://www.ebi.ac.uk/Tools/psa/emboss_needle/). For literature search and management, PubMed database (National Center for Biotechnology Information) and Citavi 4.4.0.28 (Swiss Academic Software, Wädenswil, Switzerland) were used, respectively. Protein characteristics were determined using Expasy Protparam online tool available at http://web.expasy.org/protparam/. For the calculation of solvent accessible surface areas of proteins POPS (Parameter OPtimsed Surfaces) algorithm was used (Fraternali, 2002; Cavallo, 2003). Densitometric analysis was done using ImageJ (Rasband, NIH, Bethesda, MD, USA) and EpiSearch algorithm available at http://curie.utmb.edu/episearch.html was used for mapping of conformational epitopes with phage-display (see 2.4.9) (Negi & Braun, 2009).

2.1.2 Chemicals and reagents

Unless otherwise stated, chemicals and reagents were purchased in pro analysis grade from AppliChem (Darmstadt), Carl Roth (Karlsruhe), Fermentas (St. Leon-Rot), Fluka (Neu-Ulm), Merck (Darmstadt), Milipore (Schwalbach), New England Biolabs (NEB, Frankfurt a. M.), Novagen (Darmstadt), Roche Diagnostics (Mannheim), Serva

(Heidelberg), Sigma-Aldrich (Steinheim), Qiagen (Hilden), Thermo Fisher Scientific (Schwerte) and VWR (Darmstadt).

2.1.3 Commercially available systems

Table 1: Commercially available systems and laboratory kits used.

System/Kit	Company
BSA Standard	Thermo Fisher Scientific, Schwerte
BugBuster Protein Extraction Reagent	Novagen, Darmstadt
DNA 1 kb Plus O'GeneRuler Marker	Fermentas, St. Leon-Rot
DNA 6x Orange Loading Dye	Fermentas, St. Leon-Rot
InFusion HD EcoDry Cloning Kit	Clontech, Saint-Germain-en-Laye, France
Multicolor BroadRange SDS Marker	Thermo Fisher Scientific, Schwerte
PierceTM BCA Protein Assay Kit	Thermo Fisher Scientific, Schwerte
Ponceau S Solution	Sigma-Aldrich, Steinheim
QIAprep Spin Miniprep Kit	Qiagen, Hilden
QIAquick Gel Extraction Kit	Qiagen, Hilden
QIAquick PCR Purification Kit	Qiagen, Hilden
QuikChange Lightning Multi Site-Directed Mutagenesis Kit	Agilent Technologies, Santa Clara, CA, USA
SOC medium	Sigma-Aldrich, Steinheim

2.1.4 Sera and antibodies

In total, sera of 47 patients recruited in Switzerland (Prof. Ballmer-Weber, Zürich) and Germany (PD Dr. Kleine-Tebbe, Berlin and Prof. Treudler, Leipzig) were analyzed in this project. Inclusion criteria were sensitizations to Bet v 1 and Gly m 4 and no birch pollen-specific immunotherapy. Allergy against soybean was determined by double-blind placebo-controlled food challenge (DBPCFC) or convincing clinical history. 27 patients (group 1) exhibited an allergy against soybean, while 20 patients (group 2) showed sensitization to Bet v 1 and Gly m 4 as well, but no clinical allergy to soybean. With ImmunoCAP™ analysis (Phadia 250, Thermo Fisher Scientific,

Uppsala, Sweden, see 2.4.20) specific IgE levels against Bet v 1 (code: t215), Gly m 4 (code: f353) but also soybean allergens Gly m 5 (code: f431) and Gly m 6 (code: f432), as well as total Gly m 4-specific IgG levels (code: f353) were determined for each individual patient. In total four serum pools consisting of equal ratios of three or four patients were generated, respectively. For detailed information on patient characteristics see chapter 3.1. Serum from a non-allergic patient was used as negative control and one patient with known class I food allergy to soy was used as positive control in immunoblot with rGly m 4. Both were available at Paul-Ehrlich-Institut serum collection.

For analysis of rBet v 1a and its folding variant sera of ten subjects sensitized to Bet v 1a or birch pollen were analyzed with convincing history of pollinosis to early flowering tree pollen. These patients were recruited at the University Medical Center Mainz Department of Dermatology, Mainz, Germany, and at the Klinik für Dermatologie, Venerologie und Allergologie, University of Leipzig, Germany. With ImmunoCAP™ analysis total IgE amount, as well as birch- (common silver birch, code: t3) or Bet v 1-specific IgE level (code: t215) was determined.

Antibodies:

Mouse anti-His_6 IgG_1 unconjugated/Clone 13/45/31	Dianova, Hamburg
Mouse anti-human IgE (Fc)-HRP/Clone B3102E8	Southern Biotech, via Biozol
Mouse anti-human IgG_4 HRP/Clone C8010-S840C	Southern Biotech, via Biozol
Mouse-anti-human IgG (Fc)-HRP/Clone JDC-10	Southern Biotech, via Biozol
Rabbit anti-Bet v 1 IgG/polyclonal antibody	ALK, Hørsholm, Denmark
Goat anti-rabbit IgG HRP/Clone A0545	Southern Biotech, via Biozol

2.1.5 Plasmids

Table 2: Description of DNA vectors used.

Vector	Tag	Resistance marker
pET11a	-	Ampicillin
pET15b	N-His_6	Ampicillin
pET29b	N-His_6	Kanamycin

All DNA vectors were purchased from Novagen (Darmstadt).

2.1.6 DNA oligonucleotides

Lyophilised DNA oligonucleotides used in this study were purchased from MWG Eurofins Operon (Ebersberg) in HPSF purification grade and a synthesis scale of 0.05 µmol. Subsequent Table 3 - Table 6 list all used DNA oligonucleotides.

Table 3: Oligonucleotides used for substitutions in Δ51NCS$_{N42/P49}$ model protein.

Substitution	Sequence (bp)
K4F	5'-CAT ATG GGC GTC ACA TTT GTG ATT CAT CAT GAG-3' (33 bp)
V5T	5'-ATG GGC GTC ACA AAA ACT ATT CAT CAT GAG TTG-3' (33 bp)
I6F	5'-GGC GTC ACA AAA GTG TTT CAT CAT GAG TTG GAA G-3' (34 bp)
H7E	5'-GTC ACA AAA GTG ATT GAG CAT GAG TTG GAA GTT G-3' (34 bp)
H8D	5'-ACA AAA GTG ATT CAT GAT GAG TTG GAA GTT GCT G-3' (34 bp)
E11N	5'-ATT CAT CAT GAG TTG AAT GTT GCT GCT TCA GCT G-3' (34 bp)
D18T	5'-GCT GCT TCA GCT GAT ACT ATA TGG ACT GTT TAT AG-3' (35 bp)
T21K	5'-GCT GAT GAT ATA TGG AAA GTT TAT AGC TGG CCT GGC-3' (36 bp)
P26T	5'-ACT GTT TAT AGC TGG ACT GGC TTG GCC AAG CAT C-3' (34 bp)
G27D	5'-GTT TAT AGC TGG CCT GAG TTG GCC AAG CAT CTT CC-3' (35 bp)
G27D/A29D	5'-GTT TAT AGC TGG CCT GAG TTG GAG AAG CAT CTT CCT GAC-3' (39 bp)
K30N/D34K	5'-TGG CCT GGC TTG GCC AAT CAT CTT CCT AAA TTG CTC CCT GGC GCT-3' (45 bp)
G38D	5'-CCT GAC TTG CTC CCT GAT GCT TTT GAA AAG CTA G-3' (34 bp)
A39S	5'-GAC TTG CTC CCT GGC TCA TTT GAA AAG CTA GAA-3' (33 bp)
E41K	5'-CTC CCT GGC GCT TTT AAA AAG CTA GAA AAT ATC-3' (33 bp)
K42S	5'-CCT GGC GCT TTT GAA TCA CTA GAA AAT ATC ATC-3' (33 bp)
I47E//D49N	5'-AAG CTA GAA AAT ATC GAA GGT AAT GGA GGT CCT GGT ACC-3' (39 bp)
D49N	5'-GAA AAT ATC ATC GGT AAT GGA GGT CCT GGT ACC-3' (33 bp)
D57K	5'-GGT CCT GGT ACC ATC AAA GAT ATG ACA TTT GTA-3' (33 bp)
P62E	5'-CCT GGC GCT TTT GAA TCA CTA GAA AAT ATC ATC-3' (33 bp)
G63D	5'-ATG ACA TTT GTA CCA GAT GAA TTT CCT CAT GAA TAC-3' (36 bp)
H67K	5'-CCA GGT GAA TTT CCT AAG GAA TAC AAG GAG AAG-3' (33 bp)
I74E	5'-TAC AAG GAG AAG TTT GAA TTA GTC GAT AAT GAG-3' (33 bp)
N78E	5'-TTT ATA TTA GTC GAT GAA GAG CAT CGT CTA AAG-3' (33 bp)
K84S	5'-GAG CAT CGT CTA AAG TCA GTG CAA ATG ATT GAG-3' (33 bp)
Q86S	5'-CGT CTA AAG AAG GTG TCA ATG ATT GAG GGA GGT-3' (33 bp)
L95D/G96T	5'-GGA GGT TAT CTG GAC GAT ACA GTA ACA TAC TAC ATG-3' (36 bp)
T98E/Y99K	5'-CTG GAC TTG GGA GTA GAA AAA TAC ATG GAC ACA ATC-3' (36 bp)
M101T	5'-GGA GTA ACA TAC TAC ACT GAC ACA ATC CAT GTT G-3' (34 bp)
T103D	5'-ACA TAC TAC ATG GAC GAT ATC CAT GTT GTT CCA AC-3' (35 bp)
K11N	5'-GTT GTT CCA ACT GGT AAT GAT TCA TGT GTT ATT-3' (33 bp)
S119T	5'-TGT GTT ATT AAA TCC ACT ACT GAG TAC CAT GTG-3' (33 bp)
E121K	5'-ATT AAA TCC TCA ACT AAG TAC CAT GTG AAA CCT G-3' (34 bp)

H123E	5’-TCC TCA ACT GAG TAC GAA GTG AAA CCT GAG TTT G-3’ (34 bp)
E127D	5’-TAC CAT GTG AAA CCT GAT TTT GTC AAA ATC GTT G-3’ (34 bp)
E127D/V129E	5’-TAC CAT GTG AAA CCT GAT TTT GAA AAA ATC GTT GAA CCA C-3’ (40 bp)
V132N	5’-GAG TTT GTC AAA ATC AAT GAA GAA CTT ATC ACC-3’ (33 bp)*
E133Q	5’-TTT GTC AAA ATC GTT XXX *GAA* CTT ATC ACC ACC-3’ (33 bp)*
P134D	5’-GTC AAA ATC GTT GAA GAT CTT ATC ACC ACC GGT *TC*-3’ *(35 bp)**
L135E	5’-AAA ATC GTT GAA GAA GAA ATC ACC ACC GGT TCA TTA G-3’ (37 bp)*
I136K	5’-ATC GTT GAA GAA CTT AAA ACC ACC GGT TCA TTA GC-3’ (35 bp)*
T138K	5’-GAA GAA CTT ATC ACC AAA GGT TCA TTA GCT GCC-3’ (33 bp)*
A142D	5’-ACC ACC GGT TCA TTA GAT GCC ATG GCA GAC GCC-3’ (33 bp)*
D146K	5’-TTA GCT GCC ATG GCA AAA GCC ATC TCA AAA CTT G-3’ (34 bp)
S149E	5’-ATG GCA GAC GCC ATC GAA AAA CTT GTT CTA GAA C-3’ (34 bp)
S157D/S159Y	5’-GTT CTA GAA CAC AAA GAC AAA TAC AAC CTC GAG CAC CAC-3’ (39 bp)
P134E/P140S	5’-GTC AAA ATC GTT GAA GAA CTT ATC ACC ACC GGT TCA TTA GCT GCC ATG GC-3’ (50 bp)

Shown are forward primers with underlined triplets coding for substituted amino acids. Sequences marked with an * were used with Δ51NCS$_{P134E/P140S}$ as template with grey codons representing proline substitutions. All oligos were purchased from Eurofins MWG Operon.

Table 4: Oligonucleotides used for infusion cloning of Gly m 4 amino acid sequence.

Name	Sequence	Restriction sites
Gly m 4 f	5’- AGA GAG GCT GAA GCT GAA TTC GGT GTT TTT ACC TTT GCA G-3’	NdeI/ BamHI
Gly m 4 r	5’-G TTA GCA GCC GGA TCC TTA ATT ATA ATC CGG ATG-3’	

Forward (f) and reverse (r) oligos used with Infusion EcoDry Cloning Kit to amplify Gly m 4-coding sequence with PCR after excision out of pET11a vector with NdeI and BamHI restriction enzymes. PCR product was cloned into pET15b vector digested with same restriction enzymes.

Table 5: Oligonucleotides used for amino acid substitutions in Gly m 4.

Substitution	Sequence (bp)	Comment
E6A	5’-GGT GTT TTT ACC TTT GCA GAT GAA ATT AAT AGT C-3’ (34 bp)	-
D27A	5’-CTG GTT ACC GAT GCC GCA AAT GTT ATT CCG *GCA*-3’ (33 bp)	Based on Gly m 4$_{K32A}$
K32A	5’-AAT GTT ATT CCG AAA GCA CTG GAT AGC TTT AAA AG-3’ (35 bp)	-
D35A	5’-ATT CCG *GCA* GCA CTG GCA AGC TTT AAA *GCA* GTT-3’ (33 bp)	Based on Gly m 4$_{K32A/S39A}$
S39A	5’-CTG GAT AGC TTT AAA GCA GTT GAA AAT GTG GAA GGT-3’ (36 bp)	-
E44SA	5’-GCA GTT GAA AAT GTG AGC GGT AAT GGT GGT CCG-3’ (33 bp)	-
D60A	5’-ATT ACC TTT CTG GAA GCA GGC GAA ACC AAA TTT GTG-3’ (36 bp)	-
N77K	5’-AGC ATT GAT GAA GCC AAA CTG GGC TAC AGC TAT AGC-3’ (36 bp)	-
D92A/E95A	5’-GGT GCA GCA CTG CCG GCA ACC GCA GCA AAA ATT ACA TTT GAT-3’ (42 bp)	-

T93A/E95A	5'-GCA GCA CTG CCG GAT GCA GCA GCA *GCA* ATT ACA TTT GAT-3' (39 bp)	Based on Gly m 4_{K96A}
E95A	5'-CTG CCG GAT ACC GCA GCA AAA ATT ACA TTT GAT AG-3' (35 bp)	-
K96A	5'-CCG GAT ACC GCA GAA GCA ATT ACA TTT GAT AGC-3' (33 bp)	-
K96A	5'-CCG *GCA* ACC GCA *GCA* GCA ATT ACA TTT GAT AGC-3' (33 bp)	Based on Gly m $4_{D92A/E95A}$
N108K	5'-CTG GTT GCC GGT CCG AAA GGT GGT AGC GCA GGT-3' (33 bp)	-
K118A	5'-GGT AAA CTG ACC GTT GCA TAT GAA ACC AAA GGT G-3' (34 bp)	-
E120A	5'-CTG ACC GTT AAA TAT GCA ACC AAA GGT GAT GCA GAA CGG-3' (39 bp)	-
E126A	5'-ACC AAA GGT GAT GCA GCA CCG AAT CAG GAT GAA-3' (33 bp)	-
K133E	5'-AAT CAG GAT GAA CTG GAA ACC GGT AAA GCA AAA GC-3' (35 bp)	-
Y149A	5'-AAA GCC ATT GAA GCA GCA CTG CTG GCA CAT CCG-3' (33 bp)	-
S111P	5'-GGT CCG AAT GGT GGT CCG GCA GGT AAA CTG ACC-3' (33 bp)	-
L150P	5'-GCC ATT GAA GCA TAT CCG CTG GCA CAT CCG GAT-3' (33 bp)	-

Shown are forward primers with underlined triplets coding for substituted amino acids and codons shown in *italic* representing substitutions in template not present in original Gly m 4 sequence. Template used for each oligonucleotide is listed in comment section while – corresponds to original Gly m 4 sequence. All oligos were purchased from Eurofins MWG Operon.

Table 6: Oligonucleotides used for amino acid substitutions in Bet v 1a.

Substitution	Sequence (bp)
S112P	5'-ACA CCG GAT GGT GGT CCC ATT CTG AAA ATT AGC-3' (33 bp)
R145P	5'-GGT GAA ACC CTG CTG CCT GCA GTT GAA AGC TAT-3' (33 bp)

Shown are forward primers with underlined triplets coding for substituted amino acids. All oligos were purchased from Eurofins MWG Operon.

2.2 Microbiological techniques

2.2.1 *E. coli* strains and media

E. coli strains:

Table 7: *E. coli* cloning and expression strains.

Strain	Description
BL21-CodonPlus (DE3)	F^- ompT hsdS(r_B^- m_B^-) dcm^+ Tet^r gal λ(DE3) endA Hte [argU proL Cam^r] [argU ileY leuW Strep/$Spec^r$
BL21-CodonPlus (DE3)-RIPL	F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G λ-
OneShot TOP10 chemically competent	F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G λ-
Stellar	F^-, *end*A1, *sup*E44, *thi*-1, *rec*A1, *rel*A1, *gyr*A96, *pho*A, Φ80d *lac*ZΔ M15, Δ (*lac*ZYA - *arg*F) U169, Δ (*mrr* - *hsd*RMS - *mcr*BC), Δ*mcr*A, λ–

E. coli strains were purchased from Stratagene (via Agilent Technologies, Santa Clara, CA, USA), Invitrogen (via Thermo Fisher Scientific) and Clontech (Mountain View, CA, USA).

E. coli media:

LB medium	10 g/l Tryptone EZMix™ (Sigma-Aldrich) 10 g/l Yeast extract EZMix™ (Sigma-Aldrich) 5 g/l NaCl, pH 7.0
LB agar	LB medium Agar-Agar 15 g/l antibiotics added in appropriate concentration

2.2.2 Cultivation and storage of *E. coli*

E. coli strains were grown in LB medium by shaking at 210 rpm at 37°C. Optical densities were determined by measuring absorbance at a wavelength of 600 nm. For selection of bacterial colonies antibiotics with a final concentration of 34 µg/ml chloramphenicol (Sigma Aldrich), 20 µg/ml kanamycin (Carl Roth) or 50 µg/ml carbenicillin (Carl Roth) were added to cultures and LB agar plates. For long-term storage cultures were supplemented with 10% (v/v) glycerol, snap frozen and stored at -80°C.

2.2.3 Transformation of plasmid DNA into chemically competent *E. coli*

Competent *E. coli* cells were thawed on ice. For chemical transformation 20 µl bacteria suspension was mixed with 1 µl DpnI-digested DNA from mutagenesis PCR or 0.5 µl purified plasmid DNA and incubated on ice for 30 minutes. After heat shock for 30 seconds at 42°C cells were incubated on ice for 2 minutes. Subsequently 270 µl SOC medium was added and incubated for 1 h at 37°C and 300 rpm in a thermal shaker (ThermoStat plus, Eppendorf, Hamburg). Transformed cells were selected by streaking out 100 µl of bacterial suspension on LB agar plates containing respective antibiotic(s) and incubation overnight at 37°C.

2.2.4 Expression of proteins in *E. coli*

For expression of recombinant protein from pET15b or pET29b expression plasmid, *E. coli* strain BL21-CodonPlus(DE3) or -RIPL cells were used, respectively. Open reading frame of Bet v 1a and Gly m 4 optimized for codon usage in *E. coli* was purchased from Genart (Life Technologies, via Thermo Fisher Scientific). An overnight culture of a colony was resuspended with LB and respective antibiotic(s) to 1 L and a final OD_{600} of 0.2. The cells were grown at 37°C and 210 rpm. Protein expression was induced with a final concentration of 1 mM IPTG when OD_{600} reaches 0.6-0.7. After shaking for 3 h at 37°C and 210 rpm cells were harvested by centrifugation (10,000 g, 15 min, 4 °C). Cell pellets were washed in dH_2O and centrifuged (10,000 g, 15 min, 4°C) again. Pellets were either processed directly or stored at -80°C after snap freezing.

2.3 Molecular biology

2.3.1 Preparation of chemically competent *E. coli*

RF1 buffer	30 mM Potassium acetate 10 mM $CaCl_2$ 50 mM $MnCl_2$ 100 mM RbCl 15% Glycerol pH adjusted to 5.8
RF2 buffer	100 mM MOPS 75 mM $CaCl_2$ 10 mM RbCl 15% Glycerol pH adjusted to 6.5

For generation of chemically competent bacteria 250 ml LB medium was inoculated with an overnight culture from a single colony of bacterial strain. Cells were grown at 37°C and 350 rpm until OD_{600} reaches 0.5. All following steps were performed at 4°C or with pre-cooled materials and buffers. After centrifugation at 4°C and 5,000 g for 10 min sedimented cells were resuspended in 100 ml cold RF1 buffer and stored on ice for 5 min. Pelleted cells after second centrifugation at 4°C and 4,500 g for 10 min were resuspended in 10 ml cold RF2 buffer followed by incubation on ice for 60 min. Afterwards cells were aliquoted, snap-frozen and stored at -80°C.

2.3.2 Plasmid isolation

An overnight culture (4 ml, 37°C, 210 rpm) of a cell colony in LB medium was harvested by centrifugation (10,000 g, 3 min, 4°C). Plasmid DNA was isolated from bacterial pellet using QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions.

2.3.3 Determination of DNA concentration in solution

DNA concentrations were determined with UV-Vis spectroscopy using a Nanodrop spectrophotometer (ND-1000, peQlab, via VWR) by measuring the absorbance at

wavelengths of 260 and 280 nm. An OD_{260} of 1 equals a concentration of 50 µg/ml double-stranded DNA.

2.3.4 Restriction

pET vectors were digested with BamHI and NdeI fast digest enzymes (Thermo Fisher Scientific) according to manufacturer's instructions.

2.3.5 Agarose gel electrophoreses

TAE buffer	40 mM Tris 20 mM Acetic acid 1 mM EDTA

To verify plasmid isolation or plasmid DNA restriction 1% agarose gels in TAE buffer were used. Gels were prepared and run in RGX-60 systems (Cascade Biologics, via Thermo Fisher Scientific) for 10 minutes at 210 V. DNA samples were mixed with DNA 6x Orange Loading Dye (Fermentas) and loaded onto the gel together with 2 µl of DNA 1 kb Plus O'GeneRuler Marker (Fermentas). Afterwards gels were stained for 10 minutes with ethidium bromide and visualized with an Intas Gel Jet Imager 2000 (Intas Science Imaging Instruments, Göttingen).

2.3.6 Isolation of DNA from agarose gel

After agarose gel electrophoresis DNA fragments were isolated on a UV-table (Intas Science Imaging Instruments) using a scalpel. DNA was extracted from agarose with QIAquick Gel Extraction Kit (Qiagen) according to manufacturer instructions.

2.3.7 Cloning of DNA fragments into pET15b vector

The open reading frame of Gly m 4 and Bet v 1a optimized for codon usage in *E. coli* was purchased from Geneart (Life Technologies, via Thermo Fisher Scientific) and inserted into pET11a vector. For cloning of linearized DNA fragments into bacterial

expression vector pET15b with N-terminal His-tag, In-Fusion HD EcoDry Cloning Kit (Clontech) was used according to manufacturer's instructions. Briefly, pET15b vector with inserted DNA sequence coding for Bet v 1a was digested with BamHI and NdeI to remove Bet v 1a-coding sequence from open reading frame and obtain linearized pET15b vector. Both restriction enzymes were furthermore used to linearize pET11a vector and obtain DNA sequence coding for Gly m 4. Using gene-specific oligonucleotides (see Table 4) Gly m 4-coding sequence was PCR-amplified creating 15 bp overlaps homologous to the ends of linearized pET15b vector. 50 ng of linearized pET15b vector and 50 ng of amplified Gly m 4-coding DNA fragment were incubated with In-Fusion HD EcoDry pellet for 15 min at 37°C and 15 min at 50°C. For transformation in *E. coli* stellar cells 2.5 µl of mixture was used. Successful cloning of Gly m 4-coding fragments into pET15b vector was verified by DNA sequencing (MWG Eurofins Operon).

2.3.8 Cloning of Δ51NCS$_{N42/P49}$ and Δ51NCS$_{N42/P49}$ variants

Δ51NCS$_{N42/P49}$ was generated by Jasmin Nürnberg within her Master thesis in 2014 at the Division of Allergology at the Paul-Ehrlich-Institut. Briefly, pET29a-Δ29NCS described by Berkner *et al.* (Berkner *et al.*, 2007) was truncated by eleven amino acids at both N- and C-termini and two insertions/substitutions namely N42 and P49 were introduced to adjust amino acid sequence to a typical Gly m 4-type allergen.

2.3.9 Sequencing

Sequencing of isolated plasmid DNA was done by bidirectional Sanger sequencing (MWG Eurofins Operon).

2.3.10 Site-directed mutagenesis of DNA

For insertion of single or multiple DNA mutations QuikChange Lightning (Multi) Site-Directed Mutagenesis Kit (Agilent Technologies) was used. Briefly, PCR was performed in a total volume of 25 µl with 100 ng of plasmid DNA, 50-100 ng of specific forward oligonucleotide primer, 1 µl dNTP mix, 0.75 µl QuikSolution and

0.5 µl QuikChange Lightning (Multi) enzyme blend in 10x QuikChange Lightning (Multi) reaction buffer. Amplification was done in a Thermomixer comfort (Eppendorf). The denaturing step was carried out for 20 s at 95°C, annealing for 30 s at 55°C and elongation at 65°C for 30 s/kb (3 min for pET15b and pET29a). Digestion of template DNA was done by incubation with 1 µl DpnI enzyme for 1 h at 37°C. For transformation into *E. coli* OneShot TOP10 cells 1 µl digested DNA was used. Introduction of specific mutations were analyzed by DNA sequencing.

2.4 Protein techniques

2.4.1 SDS-polyacrylamide gel electrophoresis

Sample buffer (5x)	10% (w/v) SDS
	0.1% (w/v) Bromphenol blue
	0.25 M Tris/HCl pH 6.8
	50% Glycerol
	7.7% (w/v) DTT
SDS running buffer (10x)	0.25 M Tris
	1.92 M Glycine
	1% (w/v) SDS

Recombinant proteins were analyzed under reducing conditions by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) according to the method described by Laemmli (Laemmli, 1970). For protein separation total bis/acrylamid levels of 5% for stacking and 15% for running gel were used according to composition shown in Table 8.

Table 8: Composition of stacking and running gel used in SDS-PAGE.

	5% Stacking gel	**15% Running gel**
Rotiphorese® NF-Acrylamide/bis-solution 30 %	1.7 ml	10 ml
Tris/HCl	1.25 ml (pH 6.8)	5 ml (pH 8.8)
10% SDS	0.1 ml	0.2 ml
10% APS	0.1 ml	0.2 ml
TEMED	0.01 ml	0.02 ml
dH_2O	6.8 ml	4.6 ml

Prior to loading, protein samples were mixed with sample buffer and denatured at 95°C for 5 min. Protein separation was performed at 100 V for 10 min and 200 V for 45 min using Mini-PROTEAN® 3 Cell equipment (BioRad, München) and SDS running buffer. As a molecular weight standard, Spectra™ Multicolor Low/Broad Range Protein Ladder (Thermo Fisher Scientific) was used.

2.4.2 Coomassie staining

Coomassie solution
35% Methanol
35% Glacial acetic acid
0.2% (w/v) Coomassie Brilliant Blue (G250)

Destaining solution
30% Methanol
10% Glacial acetic acid

After SDS-PAGE gels were washed for 5 min in dH_2O. For coomassie staining gels were incubated for 3 min in coomassie solution. To remove unspecific staining, gels were transferred to destaining solution for several hours.

2.4.3 Analysis of protein expression in small-scale

Pellet obtained from 2 ml bacterial culture (see 2.2.4) was used with BugBuster™ Protein Extraction Reagent (Merck Millipore) to analyze protein expression in soluble fraction and inclusion bodies. To resuspend cell pellet 0.25 ml BugBuster reagent with 0.25 µl Benzonase® Nuclease® (250 U/µl, Novagen) and 40 µl protease inhibitor Complete Mini EDTAfree (Roche Diagnostics) were added. Remaining steps were performed with same volume of BugBuster reagent and according to manufacturer's instructions. Samples of soluble fraction and inclusion bodies were analyzed by SDS-PAGE (see 2.4.1) for protein expression.

2.4.4 Purification of soluble protein

Lysis buffer	10 mM Potassium phosphate pH 7.4 10 mM Imidazole
Elution buffer 1	10 mM Potassium phosphate pH 7.4 0.5 M Imidazole

Bacterial pellet was resuspended in 15 ml lysis buffer with one tablet protease inhibitor EDTAfree (Roche Diagnostics) added. After stirring for 20 min at RT, DNase I (Applichem, via VWR) and lysozyme (from chicken egg white, 10,000 U/mg, Sigma-Aldrich) were added followed by stirring for 30 min at RT. Cells were homogenized in a cell disrupter (TS series, Constant Symstems Ltd., Daventry, UK) at 1,350 bar. After centrifugation for 15 minutes at 12,000 g and 4°C, lysate was filtered (0.2 µm, VWR) and incubated over night at 4°C with 3 ml Ni-beads (Ni-NTA Superflow, Qiagen) equilibrated in lysis buffer. Pelleted beads were used together with a Bioline HR glass column system (Knauer, Berlin). Target protein could be eluted by affinity chromatography on a HPLC Smartline System (Knauer) with elution buffer 1 (flow rate 0.3 ml/min, 2.5 CV). Eluted fractions were analyzed on SDS-PAGE and protein-containing fractions were pooled and concentrated with Centrifugal Filter Units (3,500 MWCO, Merck Millipore) if necessary.

2.4.5 Purification of proteins from inclusion bodies

Denaturing buffer	8 M Urea 10 mM Potassium phosphate pH 7.4 0.5 M NaCl 10 mM Imidazole
Renaturing buffer	0.5 M Urea 10 mM Potassium phosphate pH 7.4 10 mM Imidazole
Elution buffer 2	0.5 M Urea 10 mM Potassium phosphate pH 7.4 0.5 M Imidazole

First steps for purification of proteins from inclusion bodies were equal to purification of soluble proteins except homogenization at 1,900 bar. After centrifugation for 15 minutes at 12,000 g and 4°C, pellet was resuspended in 10 ml denaturing buffer and incubated over night at 4°C with 3 ml Ni-beads equilibrated in denaturing buffer. Pelleted beads were used together with a Bioline HR glass column system (Knauer) for affinity chromatography. On HPLC Smartline System (Knauer) a linear gradient from denaturing buffer to renaturing buffer was used to refold target protein bound to Ni-beads (flow rate 0.2 ml/min, gradient 2.5 CV). Afterwards protein was eluted from Ni-beads with elution buffer 2 (flow rate 0.3 ml/min). Eluted fractions were analyzed on SDS-PAGE and protein-containing fractions were pooled and concentrated with Centrifugal Filter Units (3,500 MWCO, Merck Millipore) if necessary.

2.4.6 Size exclusion chromatography and dialysis

Size exclusion buffer	10 mM Potassium phosphate pH 7.4

Concentrated protein, from soluble fraction or inclusion bodies, eluted from Ni-affinity chromatography was further purified by size exclusion chromatography (BioFox 40/1200 SEC, HPLC Smartline System, Knauer) with size exclusion buffer (flow rate 0.3 ml/min, 2 CV). Again fractions were analyzed by SDS-PAGE and afterwards concentrated (Centrifugal Filter Units 3,500 MWCO, Merck Millipore) if necessary. After dialysis (D-Tube™ Dialyzer, MWCO 3.5 kDa, Novagen) against 4 L of 10 mM KPi pH 7.4 at 4°C aliquots were prepared, snap-frozen and stored at -80°C.

2.4.7 Cleavage of His-Tag

His-Tag of proteins in pET15b vector was cleaved after elution from Ni-beads with thrombin from bovine plasma (Sigma-Aldrich). For this, roughly 8 mg of recombinant protein and 0.1-0.15 KU of thrombin were mixed and protein cleavage was analyzed by SDS-PAGE. To remove remaining thrombin enzyme and cleaved His-Tag, size exclusion chromatography and dialysis were performed, as described above (see 2.4.6).

2.4.8 Measuring protein concentration

Protein concentration could be measured with Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific) according to manufacturer's instructions or by measuring absorbance at 280 nm using Nanodrop spectrophotometer (ND-1000, peQlab, via VWR). Extinction coefficient ($M^{-1}cm^{-1}$) and molecular weight (Da) of proteins were determined using Expasy Protparam online tool available at http://web.expasy.org/protparam/.

2.4.9 Competitive immunoscreening of phage-displayed peptides

Competitive immunoscreening was done as described with some modifications (Mittag *et al.*, 2006; Subbarayal *et al.*, 2013). 5 mg of Dynabeads M-280 Tosylactivated (Invitrogen Dynal AS, Oslo, Norway) was coupled with 100 µg anti-human IgE antibody (Southern Biotech) for 20 h at RT. After blocking for 2 h at RT, beads were incubated with 500 µl serum for 3 h at 4°C and washed 5 times with washing buffer. For first round of panning 50 µl beads were incubated with 2^{11} phages from 12-mer peptide library (New England Biolabs) at RT overnight. For amplification *E. coli* ER2738 cells were infected with washed phages bound to beads for 4 h at 37°C. Phages were precipitated and titer could be determined. For second round of panning $2*10^{11}$ phages from first round were incubated with serum IgE bound to beads. After washing bound phages were eluted competitively by incubation with 25 µg rGly m 4 for 1 h at RT. Eluted phages were amplified again and titer was determined. Single plaques on LB plates containing IPTG and Xgal were purified and phage DNA sequence was determined. By translation of DNA sequence, amino acid composition of displayed peptide was determined. To increase specificity

of eluted phages several modifications were added to the established protocol. First, number of washing steps was increased to a maximum of 20 steps. Second, different washing but also blocking buffers were used. Washing buffers contained PBS with 0.05% Tween20, 0.1% BSA or both while blocking buffers with PBS and concentrations of 0.3% Tween20, 0.5% BSA or both were tested. Third, after each washing step the supernatant derived by magnetic separation was amplified and checked for elution of unspecifically bound phages by growing on IPTG/Xgal-plates. Fourth, negative selection was included, where amplified phages after competitive elution with rGly m 4 were incubated with magnetic beads only, followed by incubation with anti-human IgE-coupled beads. This way, phages binding unspecifically to beads and/or anti-human IgE were removed.

Amino acid sequence of identified phages were mapped onto Gly m 4 surface using EpiSearch algorithm (http://curie.utmb.edu/episearch.html) (Negi & Braun, 2009). EpiSearch uses composition of peptide sequences as input and compares these with the distribution of surface-exposed amino acid patches on the allergen's three-dimensional structure. Results are expressed as score value for each individual patch on allergen surface. For mapping, default settings for 12-mer peptides were used and results with scores ≥1 were considered as positive putative epitopes. Three-dimensional structure of Gly m 4 with pdb code 2K7H was used as template.

2.4.10 Immunoblot

Anode buffer 1	0.3 M Tris pH 10.4 20% (v/v) Methanol
Anode buffer 2	25 mM Tris pH 10.4 20% (v/v) Methanol
Cathode buffer	40 mM 6-Aminohexanoic acid pH 7.6 20% (v/v) Methanol
Blocking buffer	TBS with 0.3% Tween20™
Washing buffer	TBS with 0.05% Tween20™
Incubation buffer	TBS with 0.05% Tween20™, 0.1% BSA

Proteins separated after SDS-PAGE (see 2.4.1) were transferred to a nitrocellulose membrane (0.2 µm, Whatman, via Sigma-Aldrich) using semi-dry blotting. Nitrocellulose membrane was equilibrated for 30 min in anode buffer 2, while filter papers (GB003, Whatman) were briefly dampen with anode buffer 1, anode buffer 2 or cathode buffer, respectively. Protein transfer occurred for 1 h at constant current of 0.8 mA per cm^2 of gel. To verify quantitative blotting, transferred proteins could be stained with 0.1% (w/v) Ponceau S Solution in 5% acetic acid (Sigma-Aldrich). Afterwards membrane was blocked for unspecific binding with blocking buffer for 1 h at RT and cut into strips of 0.3 cm width each. After incubation for 10 min in washing buffer 120 µl of primary antibody or serum diluted 1:10 in incubation buffer were added for at least 1 h. Membrane was washed 5 times with washing buffer for 5 min each. Secondary HRP-coupled mouse anti-human IgE antibody (Southern Biotech, via Biozol, Eching) diluted 1:10,000 in incubation buffer was added to the membrane for 1 h. After washing chemiluminescence of horseradish peroxidase (HRP)-conjugated antibodies was detected using LumiGLO Reserve™ Chemiluminescent Substrate Kit (KPL, Gaithersburg, MD, USA) and Imager system Fusion FX (Vilber Lourmat, Eberhardzell) or Curix 60 processor (AGFA, Leverkusen) and Hyperfilm MP (Amersham, via GE Healthcare, Freiburg) together with Hypercassette Autoradiography Cassettes (Amersham, via GE Healthcare). In case secondary antibody with AP conjugation was used staining was performed with BCIP (5-bromo-4-chloro-3'-indolyphosphate) and NBT (nitro-blue tetrazolium) solution for 30 s. Staining was stopped with dH_2O and membrane was scanned. For detection of IgG binding to rBet v 1a and rBet v $1a_{S112P/R145P}$ variant on nitrocellulose membranes, Bet v 1-specific polyclonal rabbit IgG (ALK) was diluted 1:10,000 in incubation buffer and added for 1 h. Bound IgG was detected with goat-anti-rabbit IgG conjugated with horseradish peroxidase (Southern Biotech, via Biozol) diluted 1:10,000 in incubation buffer as described above.

For parallel detection of IgE binding different strips were incubated in the same detection solution containing serum diluted 1:10 in incubation buffer. Each strip was derived from an immunoblot with a different antigen, competing for IgE present in serum samples. Residual steps of immunoblotting were carried out as described above.

2.4.11 Dot blot

Dot blot analysis of proteins was carried out according to immunoblotting described in chapter 2.4.10, but without prior SDS-PAGE and semi-dry blotting. In contrast 1 µg of each recombinant protein was spotted directly on nitrocellulose membrane. After 15 min air-dried membranes could be stained with Ponceau S solution to check for presence of membrane-bound protein. Afterwards membrane was blocked and incubated with 50 µl 1:10 diluted serum as described for immunoblot. Chemiluminescence was detected with 1:10,000 diluted secondary HRP-coupled mouse anti-human IgE antibody (Southern Biotech, via Biozol) in incubation buffer (see 2.4.10), LumiGLO Kit and Imager system. Analysis of densitometry was evaluated using ImageJ (Rasband, NIH, Bethesda, MD, USA).

2.4.12 Enzyme-linked immunosorbent assay (ELISA)

Coating buffer	PBS
Washing buffer	PBS with 0.05% Tween20™
Blocking buffer	PBS, 2% BSA
Incubation buffer	PBS with 0.05% Tween20™, 0.1% BSA
Substrate solution	210 mM Potassium citrate pH 3.9 1 mM TMB 0.01% H_2O_2

With Enzyme-linked immunosorbent assay (ELISA) binding between antibodies and antigens can be analyzed on solid phase. 50-200 ng of recombinant allergen was adsorbed over night at RT to microtiter plates (Maxisorp, Nunc, Wiesbaden) in a total volume of 100 µl/well coating buffer. Plates were washed four times with 300 µl washing buffer. After blocking for 2 h at RT with 100 µl/well blocking buffer, primary antibody or serum diluted in incubation buffer was added. Plate was washed on ELISA washer and incubated for 1 h with 100 µl/well secondary HRP-coupled mouse anti-human IgE antibody (Southern Biotech, via Biozol) diluted 1:1,000 in incubation buffer. For visualization 100 µl/well 3,3′,5,5′-tetramethylbenzidine (Roth) was used as substrate for the horseradish peroxidase and the absorbance at 450 nm was measured after stopping the reaction after 10 min with 25% H_2SO_4 on

SpectraMax 250 Microplate Reader (Molecular Devices, Biberach an der Riss). Non-specific binding of secondary antibody was analyzed in control reaction without adding primary antibody or serum, respectively.

2.4.13 ELISA-Inhibition

ELISA-Inhibition was performed according to ELISA method described in 2.4.12 with slight modifications. After coating and blocking, dilutions of serum and inhibitor pre-incubated for 1 h were added. In general one dilution of serum was mixed with several concentrations of allergen inhibitor to analyze inhibition reaction between serum IgE, coated allergen and inhibitor. After incubation for 3 h at RT plates were washed and protocol follows the one described above. Control reactions contained unspecific binding of secondary antibody where no serum and inhibitor were added, as well as maximal binding control of serum IgE without adding allergen inhibitor.

2.4.14 IgE depletion experiments

For measuring inhibition of IgE binding between NCS variants and rGly m 4 IgE depletion experiments were used. 200 ng/well rGly m 4 were coated over night at 4°C in microtiter plates (Maxisorp, Nunc) in PBS. Plates were washed and blocked according to section 2.4.12. In parallel, different amounts of rGly m 4, $\Delta NCS_{N42/P49}$, $\Delta NCS_{N42/P49}$ variants or BSA were incubated with 300 µl 1:10 diluted serum for 3 h followed by precipitation of allergen-antibody complexes with Ni-beads (Ni-NTA Superflow, Qiagen). Supernatant containing unbound IgE antibodies was removed and added to blocked wells of ELISA plates. After incubation for 1.5 h plates were washed and bound IgE was measured with HRP-conjugated secondary antibody as described for ELISA (see 2.4.12). Residual IgE binding of supernatants to plate-coated rGly m 4 was calculated by comparing IgE binding to rGly m 4 upon depletion with control protein BSA set to maximum IgE binding (100%).

2.4.15 Mediator release from humanized Rat Basophilic Leukaemia (RBL) cells

Mediator release assay was performed according to protocol established by Vogel *et al.* (Vogel *et al.*, 2005). Briefly, transfected RBL-2H3 cells expressing α-chain of

human FcεRI were passively sensitized with serum IgE antibodies (dilution 1:40). After washing of the cells, addition of different concentrations of allergen or allergen mixtures induced cross-linking of receptor-bound IgE antibodies causing cell degranulation. Degranulation was quantified by photometric measurement of β-hexosaminidase in culture supernatants. Percentage of released β-hexosaminidase was calculated compared to total degranulation of cells lysed with Triton X-100 (Sigma-Aldrich) and corrected for spontaneous release of cells incubated with serum only.

2.4.16 Circular dichroism spectroscopy

Far UV circular dichroism (CD) spectra of purified proteins were recorded using a Jasco J-810 spectropolarimeter (Japan Spectroscopic, Gross-Umstadt). Measurements were performed at 293 K and constant nitrogen flow rate in 1 mm quartz cuvettes with a band width of 1 nm and a sensitivity of 100 mdeg. Each measurement comprised the average of 10 repeated scans between 255 and 185 nm at a scanning speed of 50 nm/min. Proteins were analyzed at concentrations of 5 to 10 µM in 10 mM potassium phosphate buffer at pH 7.4. Spectrum of buffer alone was subtracted. Results are shown as mean residual ellipticity.

2.4.17 NMR spectroscopy

NMR buffer	20 mM Sodium phosphate pH 7.0 0.04% Sodium azide 10% D_2O

NMR samples were prepared in NMR buffer containing 30 µM of recombinant protein. Standard 1D-^{1}H spectra with WATERGATE solvent suppression were recorded on a Bruker Avance 700 MHz spectrometer at 298 K (Piotto *et al.*, 1992). NMR data were processed and visualized with the Bruker spectrometer software TopSpin. NMR analyses were performed by Dr. Christian Seutter von Loetzen, Dr. Kristian Schweimer and Prof. Paul Rösch at the Department of Biopolymers of the University of Bayreuth.

2.4.18 Dynamic light scattering

The hydrodynamic radii (R_H) of proteins were analyzed with dynamic light scattering (DLS). Protein samples with concentrations of at least 5 µM were used after centrifugation at 4°C and 16,000 g for 15 min in Zetasizer Nano-ZS (Malvern, Herrenberg) at 25°C and a wavelength of 633 nm. Three individual measurements with 10 runs each were carried out in UV micro cuvettes (Carl Roth) with a layer thickness of 10 mm.

2.4.19 Mass spectrometry

Recombinant proteins were separated by SDS-PAGE, stained with coomassie, excised and subjected to liquid chromatography mass spectrometry (LC-MS) analysis. MS analyses were performed by Luisa Schwaben and Dr. Andreas Reuter at the Proteomics Core Facility of the Paul-Ehrlich-Institut. The experimental procedures will be described to an extent that allows the reader to assess the validity of the results.

In gel digestion:

Destaining solution	40% Ethanol, 50mM NH_4HCO_3
Reduction solution	65 mM DTT in 50 mM NH_4HCO_3
Alkylation solution	260 mM Iodoacetamide in 50 mM NH_4HCO_3
Elution buffer	25 mM NH_4HCO_3, 10% acetonitrile

The protein spots were excised and destained three times for 15 min in destaining solution. Reduction and alkylation of cysteine residues was carried for 15 minutes at 200 rpm (thermomixer, Eppendorf) with reduction and alkylation solution, respectively. The gel plugs were dehydrated with 100% acetonitrile, vacuum dried (Savant Speed Vac®, Thermo Fisher Scientific) and were rehydrated in 25 mM NH_4HCO_3 containing 75 ng/µl of trypsin (trypsin from porcine pancreas, Sigma-Aldrich). After initial digestion for 3 h at 37°C, elution buffer was added and further digestion was allowed over night at 37°C in a thermal cycler (Bibby Scientific, Staffordshire, UK). The digestion was stopped by adding of 5% formic acid to a

volume of 10% of the total digestion mixture. The samples were stored at -80°C until further MS analyses.

LC-MS analyses:

Trypsin digested proteins were analyzed using the nanoACQUITY® UPLC online coupled with nano ESI interface to a Q-TOF MS (Synapt, Waters, Manchester, UK). The solvent system consisted of solvent A (water with 0.1% (v/v) formic acid), and solvent B (acetonitrile with 0.1% (v/v) formic acid). Peptides were initially trapped on-line at a flow rate of 5 µl/min using the trap column (nanoACQUITY Trap C18, 5 µm particle size, 180 µm × 20 mm). The peptide mixture was separated using a reverse-phase analytical column (nanoACQUITY C18, 1.7 µm particle size, 100 µm × 100 mm; Waters) at a flow rate of 0.5 µl/min. The solvent composition was at 97% A for 1 min, followed by a linear gradient to 60% A for 30 min, and continued with 95% B for 1 min. The column was subsequently equilibrated at 97% A for 18 min. Glu-1-Fibrinopeptide at 1pmol/µl, delivered from the auxiliary pump of the nanoACQUITY® UPLC at a 0.5 µl/min flow rate to the reference sprayer of the interface was measured every 20 sec to serve as a lock mass during the entire sample run. The mass spectrometer was operated in data independent MS^E mode at a positive polarity and V mode (Silva *et al.*, 2005; Silva *et al.*, 2006). Scan rate was set to 0.4 sec. The collision energy was at a constant value of 4 V during low-energy scans and ramped from 15-30 V during high energy scans. Data were acquired from 50 to 1,990 m/z.

Data processing and database search:

Raw data files were processed using ProteinLynx Global Server (PLGS) version 2.4 (Waters). The processing parameters were set as follows: Retention time window, chromatographic peak width and MS TOF resolution were set to automatic, lock mass for charge state +1 and for charge state +2 were defined as 684.3469 Da and 785.8426 Da, respectively. Low energy threshold was set to 250.0 counts, elevated energy threshold to 100.0 counts and intensity threshold to 1,500 counts. The search parameters for protein identification were a maximum of one missed cleavage site, a minimum of three fragment ion matches per peptide, a minimum of seven fragment matches per protein and a minimum of one peptide match per protein with a set false positive rate of 4%. Carbamidomethylation of cysteine (C) was defined as a fixed

modification. Variable modifications were restricted to deamination of asparagine (N) and glutamine (Q), and oxidation of methionine (M). The identities of the recombinant proteins were confirmed searching an in-house database consisting of the UniProt database (as of May 2011, restricted to reviewed entries of eukaryotic organisms) and the amino acid sequences of recombinant Gly m 4 wild type (internal accession number PEI015; UniProt accession number P26987), recombinant Gly m $4_{S111P/L150P}$ (internal accession number PEI059), recombinant Bet v 1a (internal accession number PEI037; UniProt accession number P15494), recombinant Bet v $1a_{S112P/R145P}$ (internal accession number PEI043) and recombinant Gly m 4 11-fold mutant (internal accession number PEI014).

2.4.20 Determination of allergen-specific antibodies

Specific IgE and IgG antibodies to recombinant allergens were quantified by ImmunoCAP™ using an ImmunoCAP™ 250 instrument (Thermo Fisher Scientific, Uppsala, Sweden). ImmunoCAP™ measurements were performed by Michaela Gubesch at the Paul-Ehrlich-Institut.

3 RESULTS

3.1 Study population

Sera of 47 patients were included in the study (Table 9). All patients were sensitized to both Bet v 1 and Gly m 4. Inclusion criteria were an allergy against birch pollen and specific IgE against soybean allergen Gly m 4, as determined by ImmunoCAP™. 27 patients (group 1) exhibit an allergy against soybean as determined by double-blind placebo-controlled food challenge (DBPCFC) or convincing clinical history of severe symptoms upon consumption of soybean containing food. The symptoms after soybean consumption in group 1 ranged from mild to severe reactions emphasizing clinical relevance of Gly m 4-related soy allergy. Mild symptoms like blisters, conjunctivitis, drowsiness, gastrointestinal discomfort, hoarseness, nasal secretion, oral allergy syndrome (OAS) and rhinitis were observed. 70% of patients (19/27) showed severe reactions to soy like dyspnea, flush, face swelling, throat tightness or urticaria. Group 2 comprised 20 patients with sensitization to Bet v 1 and Gly m 4 as well, but without any allergic symptoms upon consumption of soybean containing protein as confirmed by food challenge. Due to low available volumes of several sera two serum pools were generated in group 1 and 2, respectively. Serum pool No. 40 was made out of three sera from soybean allergic patients (No. 40 a, b, c) while serum pool No. 41 consisted of four sera from soybean positive sera (No. 41 a, b, c, d). Both serum pools No. 43 and 44 were made out of four soybean negative sera (No. 43 a, b, c, d and No. 44 a, b, c, d), respectively. Some experiments in this study were performed with a pool of all patients of group 1 or group 2, respectively. Patients' characteristics and results of ImmunoCAP™ measurements are summarized in Table 9.

Table 9: Characteristics of serum donors of this study.
Allergen-specific IgE and IgG of sera of patients sensitized to Gly m 4 with (A, patients 1-41) and without (B, patients 16-50) birch-related allergy to soy and clinical symptoms after DBPCFC with soybean protein are shown.

A

		ImmunoCAP™				
		sIgE for birch and soybean allergic patients [kU_A/L]				tIgG [mg_A/L]
Patient No.	**Symptoms to soybean (DBPCFC or case history)**	**Bet v 1 (class)**	**Gly m 4 (class)**	**Gly m 5 (class)**	**Gly m 6 (class)**	**Gly m 4**
1	OAS, TT	90.4 (5)	8.64 (3)	0.02 (0)	0.03 (0)	14.5
2	OAS	>100 (6)	46.3 (4)	0.08 (0)	0.08 (0)	7.3
3	OAS	67.8 (5)	14.5 (3)	0.00 (0)	0.01 (0)	27.6
4	OAS, F, C	>100 (6)	>100 (6)	0.47 (1)	0.77 (2)	13.6
5	OAS, RH, TT, H	33.5 (4)	14.2 (3)	0.01 (0)	0.01 (0)	5.4
7	OAS, TT, H	28.3 (4)	8.69 (3)	0.00 (0)	0.00 (0)	3.9
8	F, OAS, B	31.3 (4)	9.84 (3)	0.01 (0)	0.02 (0)	2.9
9	OAS	14.6 (3)	3.93 (3)	0.00 (0)	0.00 (0)	5.2
10	n.d.	7.48 (3)	3.34 (2)	0.00 (0)	0.00 (0)	6.6
11	OAS	84.1 (5)	22.6 (4)	0.00 (0)	0.01 (0)	36.9
13	OAS	>100 (6)	23.7 (4)	0.06 (0)	0.08 (0)	16.0
22	OAS	52.9 (5)	3.85 (3)	0.00 (0)	0.02 (0)	8.2
28	DY, HU	42.7 (4)	9.98 (3)	0.00 (0)	0.00 (0)	8.0
30	FS, DY, HU	48.8 (4)	11.4 (3)	0.01 (0)	0.02 (0)	6.5
33	FS, OAS, TT	8.02 (3)	4.00 (3)	0.00 (0)	0.00 (0)	8.2
34	GD	0.01 (0)	0.00 (0)	0.00 (0)	0.01 (0)	6.6
35	FS, OAS, DY, DR, GD, NS	25.3 (4)	7.09 (3)	0.00 (0)	0.00 (0)	9.7
36	FS, OAS, TT, HU, DR	33.9 (4)	11.2 (3)	0.00 (0)	0.01 (0)	7.8
37	FS, OAS, TT, HU, GD	13.8 (3)	2.68 (2)	0.00 (0)	0.00 (0)	35.8
38	FS, OAS, TT	58.1 (5)	22.9 (4)	0.03 (0)	0.02 (0)	5.7
40	Serum pool, 3 patients	40.1 (4)	7.76 (3)	0.01 (0)	0.02 (0)	13.1
a	FS, OAS, TT, DR, GD					
b	FS, OAS, TT					
c	FS					
41	Serum pool, 4 patients	22.8 (4)	5.95 (3)	0.02 (0)	0.02 (0)	9.9
a	FS, OAS, HU, DR					
b	FS, OAS, TT					
c	FS, OAS, TT, DY					
d	FS, OAS, TT					

B

		ImmunoCAP™				
		sIgE for birch allergic patients without soybean allergy [kU_A/L]				tIgG [mg_A/L]
Patient No.	**Symptoms to soybean (DBPCFC)**	**Bet v 1 (class)**	**Gly m 4 (class)**	**Gly m 5 (class)**	**Gly m 6 (class)**	**Gly m 4**
16	-	49.5 (4)	7.18 (3)	0.00 (0)	0.00 (0)	36.9
20	-	28.5 (4)	5.42 (3)	0.01 (0)	0.03 (0)	26.5

24	-	10.9 (3)	3.66 (3)	0.00 (0)	0.00 (0)	15.3
25	-	7.64 (3)	1.22 (2)	0.18 (0)	0.21 (0)	61.3
26	-	>100 (6)	>100 (6)	0.08 (0)	0.10 (0)	14.7
27	-	>100 (6)	55.9 (5)	0.16 (0)	0.24 (0)	6.1
43	- Serum pool, 4 patients	30.6 (4)	6.04 (3)	0.17 (0)	0.26 (0)	9.1
44	- Serum pool, 4 patients	18.3 (4)	0.65 (1)	0.07 (0)	0.14 (0)	6.6
45	-	6.07 (3)	1.32 (2)	0.00 (0)	0.03 (0)	7.5
46	-	8.89 (3)	0.94 (2)	0.00 (0)	0.00 (0)	4.5
47	-	47.4 (4)	17.7 (4)	0.00 (0)	0.00 (0)	7.3
48	-	0.00 (0)	0.00 (0)	0.00 (0)	0.00 (0)	4.0
49	-	>100 (6)	3.31 (2)	0.04 (0)	0.02 (0)	n.d.
50	-	>100 (6)	21.1 (4)	0.04 (0)	0.02 (0)	n.d.

Sera No. 40, 41, 43 and 44 are pools of three (No. 40) or four (No. 41, 43, 44) individual soybean positive or negative sera, respectively. Specific IgE levels were measured in kU_A/L with resulting CAP class shown in parentheses. Gly m 4-specific total IgG levels were measured with a 1:10 dilution of sera in mg_A/L with a Phadia250 ImmunoCAP™. Abbreviations of clinical symptoms: B – blisters; C – conjunctivitis; DR – drowsiness, vascular dysregulation; DY – dyspnea; F – flush; FS – face swelling; GD – gastrointestinal discomfort, vomiting; H – hoarseness; HU – hives, urticaria; NS – nasal secretion, nasal congestion; OAS – oral allergy syndrome; RH – rhinitis; TT – throat tightness, swallowing discomfort; n.d. – not determined.

Specific IgE (sIgE) to Bet v 1 and Gly m 4 in the sera of the study population were quantified by ImmunoCAP™. Overall 94% of sera (34/36) had increased (CAP class ≥ 3) allergen-specific IgE levels to Bet v 1. A slightly increased median level of Bet v 1-specific IgE of 37.0 kU_A/L for group 1 (0.01 to >100 kU_A/L) compared to 29.6 kU_A/L (0 to >100 kU_A/L) in patients merely sensitized but without allergy to soy (group 2) was determined. Furthermore almost all sera (34/36, 94%) had specific IgE antibodies against Gly m 4 indicating a sensitization to the Bet v 1-homologous soybean allergen. Gly m 4-specific median IgE levels of 9.3 kU_A/L (0 to >100 kU_A/L) in group 1 and 4.5 kU_A/L (0 to >100 kU_A/L) in patients without soybean allergy (group 2) were measured. In both groups of the study population 22 (group 1) and 14 (group 2) sera with CAP classes from 0 to 6 for sIgE against Gly m 4 were present. For two patients (No. 34 and 48) no sIgE antibodies against Bet v 1 and Gly m 4 could be detected. Serum of patient 34 has already been used in Kleine-Tebbe *et al.* in 2002 (Kleine-Tebbe *et al.*, 2002) as serum of a subject sensitized to Bet v 1 and with symptoms after soybean consumption. In addition patient 48 had no IgE against Gly m 4 or Bet v 1 but had been clinically characterized as allergic to soy. To exclude a class I soybean allergy 36 sera of group 1 and 2 were analyzed for specific IgE levels against soybean allergens Gly m 5 and Gly m 6. In total 95.5% (21/22) of patients in group 1 and 100% (14/14) of patients in group 2 had very low amounts of

sIgE (CAP class 0) against both Gly m 5 and Gly m 6. Only patient 4 had CAP class 1 for Gly m 5 and CAP class 2 for Gly m 6, respectively, indicating a possible class I food allergy to soy. However very high serum levels of IgE against Bet v 1 and Gly m 4 (>100 kU_A/L) in this patient suggest a birch-related rather than a classical soybean allergy. Analyzing the total Gly m 4-specific IgG levels of group 1 and group 2, a comparable median amount of 8.1 mg_A/L and 8.3 mg_A/L of IgG was measured, respectively. Gly m 4-specific IgG levels range from 2.9 to 36.9 mg_A/L for group 1 and 4.0 to 61.3 mg_A/L for group 2.

To further analyze Gly m 4 and soybean-specific IgE levels sera of group 1 and 2 were analyzed in immunoblot (Figure 2). IgE binding of each serum to 0.33 µg rGly m 4 and 6.7 µg soybean extract was detected at different exposure times to compare IgE signal intensities derived from rGly m 4 and soy extract, respectively.

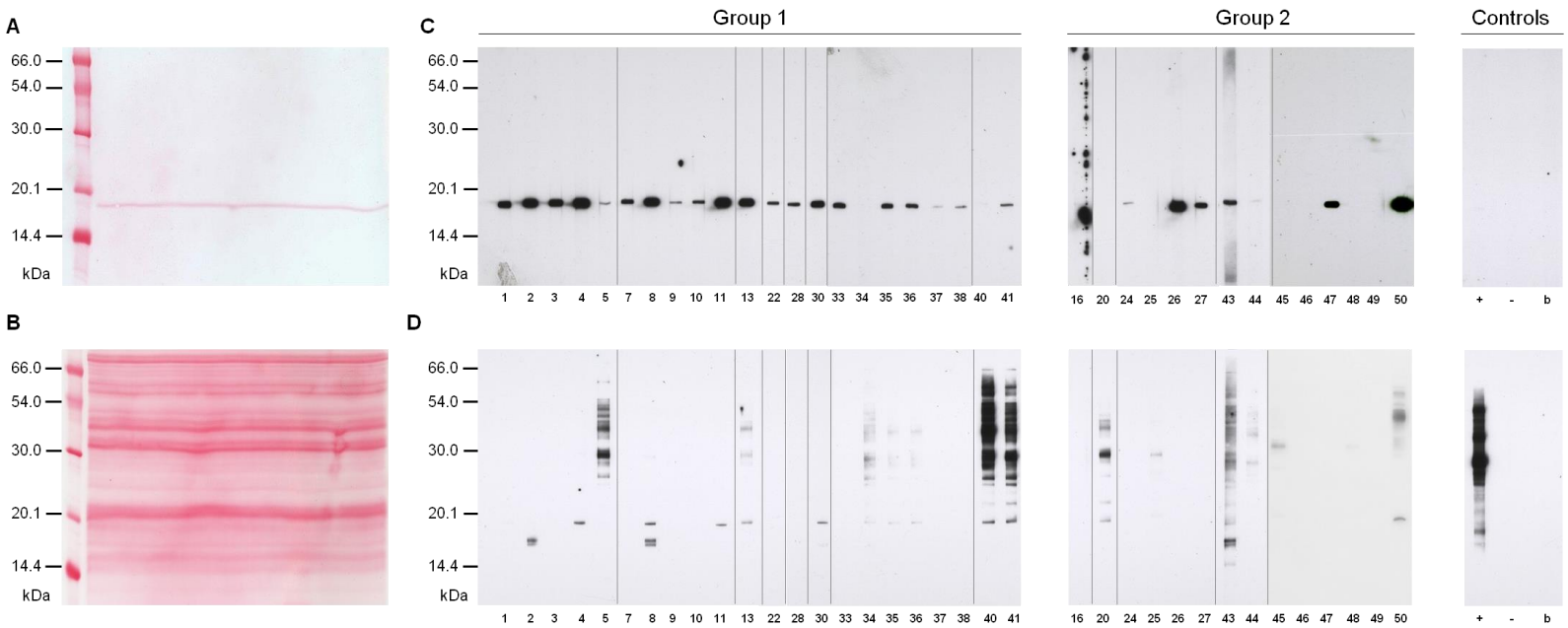


Figure 2: Binding of serum IgE to rGly m 4 and soybean extract.

Sera of patients sensitized to Gly m 4 with (group 1) and without (group 2) allergy to soy were analyzed in IgE immunoblot. Ponceau S stainings of immunoblots with rGly m 4 (A) and soybean extract (B) are shown. In total, 6.7 µg of rGly m 4 and 134 µg of soybean extract were blotted. (C) Detection of IgE binding to rGly m 4 after an exposure of 1 min. (D) Detection of IgE-binding proteins in soybean extract after 15 s of exposure. For both blots patients of group 1 and 2 were analyzed with numbers indicated. Each serum was used in 1:12 dilution and incubated over night with 0.33 µg rGly m 4 or 6.7 µg soybean extract (total protein). Serum of a patient with clinically confirmed soybean allergy (+) and serum of a non-atopic individual (-), both available from Paul-Ehrlich-Institut serum collection, as well as buffer (b) were used as controls.

With Ponceau S one transferred protein band at 17 kDa representing recombinant Gly m 4 was stained (A). In contrast analysis with soybean extract resulted in several bands at different molecular weights representing numerous proteins present in soy extract (B). According to inclusion criteria and the results of ImmunoCAP™ analyses all patients had Gly m 4-specific IgE antibodies. In immunoblot, however, IgE binding to rGly m 4 could be detected only in 95% (21/22) of sera in group 1 and 64% (9/14) of sera in group 2. In total, in sera of six patients no IgE binding to rGly m 4 could be detected in immunoblot analysis. These were serum of patient 34 for group 1 and sera of patients 25, 45, 46, 48 and 49 for group 2. These patients had no (No. 34 and 48) or low levels (No. 25, 45, 46 and 49) of Gly m 4-specific IgE according to ImmunoCAP™ and might be below detection limit for IgE antibodies with immunoblotting and used serum dilution.

Soybean extract contained several IgE-binding proteins. In immunoblot analysis sera of several patients, for example patients 5, 40, 41 and 43 had IgE binding to numerous of these soybean proteins especially with higher molecular weight compared to Gly m 4 (17 kDa). IgE signals in the range of 50-70 kDa might be related to monomers of soy seed storage proteins Gly m 5 and Gly m 6 with molecular weights of the native oligomers of 140-180 kDa and 300-380 kDa, respectively. Furthermore presence of antibodies directed against cross-reactive carbohydrate determinants (CCD), carbohydrate moieties of allergenic molecules, which might be able to induce allergic response, are possible. Some patients with sIgE against Gly m 4 showed no IgE binding at 17 kDa in immunoblot with soybean extract probably due to the very low amount of Gly m 4 in soybean extracts which might be under the detection limit for some of the sera tested. Patient 34 had IgE against several proteins of the soybean extract including a weak signal at a molecular weight of about 17 kDa comparable to Gly m 4. This is consistent with the results of ImmunoCAP™ analysis where no sIgE against Gly m 4 as well as Bet v 1 was found.

3.2 Preliminary epitope profile of Gly m 4

To identify amino acids of Gly m 4 that are involved in immunoglobulin E binding a combination of experimental and theoretical approaches was carried out. Thus, data on potential epitopes from screening of a phage-displayed peptide library for Gly m 4-specific IgE-binding, bioinformatical approaches as well as published data on

putative antibody-binding sites of allergens homologous to Gly m 4 were considered to generate a preliminary epitope profile of Gly m 4 that was to be analyzed experimentally in more detail.

3.2.1 Mapping of putative IgE-binding peptides onto Gly m 4 surface

To identify peptides that bind Gly m 4-specific IgE of serum from subjects with birch-related soy allergy I made use of an already published protocol using a phage-displayed peptide library (Mittag *et al.*, 2006). In this approach phage-displayed peptides are bound by immobilized serum IgE and competitively eluted by recombinant Gly m 4 (see 2.4.9). Eluted phages are amplified and the amino acid composition of the displayed dodecapeptide is determined by DNA sequencing and translation of the peptide-encoding DNA sequence. The peptide sequence is then mapped onto the protein surface of Gly m 4 with the web server EpiSearch to localize putative IgE-binding surface areas of the allergen (Negi & Braun, 2009).

In this study phage-display was performed with two rounds of panning as described (see 2.4.9). Despite of several modifications in phage-display protocol the number of false-positives could not be reduced. Furthermore about 20% of false-positives were also observed when buffer or a non-atopic serum control instead of a serum containing Gly m 4-specific IgE antibodies was used. This indicates an unspecific interaction of phages with any of the materials used in experimental setup. Possible unspecific binding of phages might occur with magnetic beads or anti-IgE antibodies. According to this, the phage-display approach was no longer pursued. Instead i) already published mimotope data by Mittag *et al.* where 21 peptides for soybean-allergic patients were identified and re-analyzed as described above (see Table A 1 in the appendix) (Mittag *et al.*, 2006) and ii) an alternative bioinformatical approach was used in which 7-mer peptides representing the total amino acid sequence of Gly m 4 were chosen and mapped with an off-set of one amino acid as potential IgE-binding mimotopes onto the molecular surface of Gly m 4 using EpiSearch as described (see 2.4.9). Both approaches use peptides as input, either identified experimentally via phage-display or based on amino acid sequence, followed by mapping of peptide sequences onto molecular surface of Gly m 4 to identify putative conformational epitopes. This is in line with the mechanism of an allergic response where the food allergen is digested into small peptides during gastrointestinal uptake.

These peptides are then presented e. g. on the surface of dendritic cells interacting with T cells resulting finally in an allergic response to the respective food. The analysis resulted in a total number of 152 peptides representing 270 potential epitopes distributed over 39 patches on the molecular surface of Gly m 4 (see Table A 2 in the appendix). For both analyses those predicted epitopes that were largely overlapping were combined and assigned a particular potential epitope with a specific center amino acid (Table 10 and Figure 3). A detailed overview of all mimotopes and their corresponding residues identified with theoretical and experimental mapping is shown in Table A 1 and Table A 2 in the appendix.

Table 10: Predicted epitopes of Gly m 4.

Gly m 4 peptides			Phage-displayed peptides		
Center residue	**No. of mapped peptides**	**% of total mapped peptides**	**Center residue**	**No. of mapped peptides**	**% of total mapped peptides**
94	45	30	94	6	29
119	37	24	119	6	29
29	19	19	37	4	19
37	19	13	30	2	10
150	14	9	29	1	5
30	12	8	150	1	5

Shown are six putative epitopes with their respective center residues obtained upon bioinformatic mapping of peptide sequences onto the molecular surface of Gly m 4. For theoretical and experimental approach a total number of 152 Gly m 4-specific peptides and 21 phage-displayed peptides were mapped using the EpiSearch algorithm. Experimentally determined IgE-binding peptide sequences were taken from Mittag *et al.* 2006.

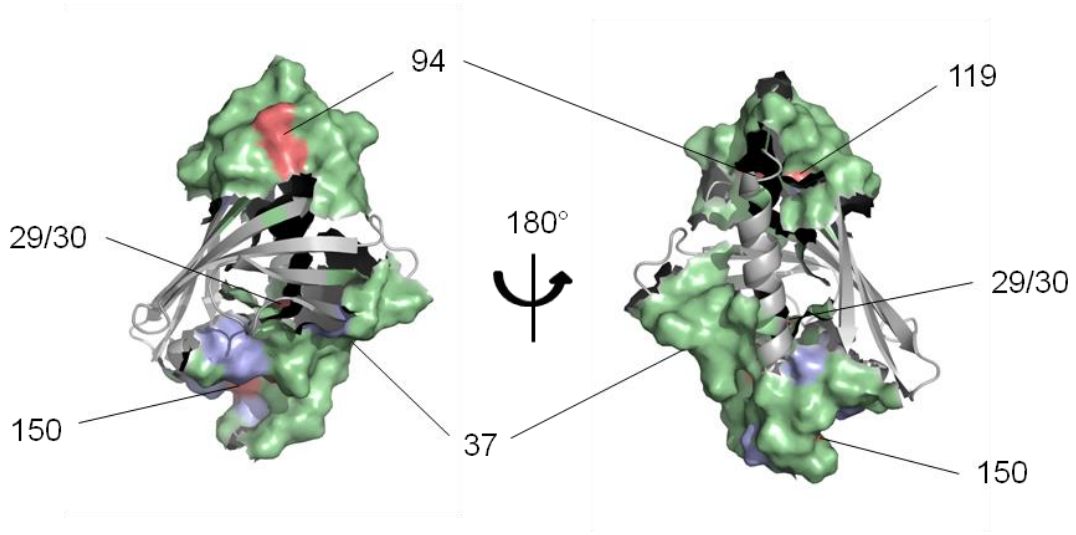


Figure 3: Comparison of predicted epitopes between theoretical and experimental mapping on Gly m 4.

Potential epitopes of Gly m 4 with respective center residues (red) identified most frequently in the bioinformatical approach using the EpiSearch algorithm. Amino acids shown in green were identified with both, while blue amino acids were predicted by Gly m 4-specific peptides only.

Interestingly putative epitopes identified most frequently in the bioinformatical approach were also identified with the experimentally observed peptides by Mittag *et al.* (Table 10) (Mittag *et al.*, 2006). Center residues 94 and 119 of putative epitopes were identified with 30% and 24% in the theoretical and with 29% each of the total number of peptides analyzed in the experimental approach. Four center residues resulting from experimental approach were identified with more than one of the total 21 eluted peptides. These center residues are 30, 37, 94 and 119 and were also identified with high frequency in theoretical approach. Furthermore both approaches showed comparable frequencies of mapped peptides for these center residues with ratios experimental/theoretical of 0.97 (29%/30%), 1.21 (29%/24%), 1.46 (19%13%) and 1.25 (10%/8%) for 94, 119, 37 and 30, respectively. Two mimotopes with center residues 29 and 150 were further mapped with more than one peptide using theoretical approach. Both were also identified as mimotopes with phage-display each comprising one mapped peptide. Ratios experimental/theoretical are 0.26 (5%/19%) and 0.56 (5%/9%) for center residues 29 and 150, respectively. Each predicted epitope consists of 8-17 amino acids with a total number of 69 different amino acids identified as putative relevant in IgE binding. Interestingly, overall results concerning mimotopes and respective center residues are comparable in both approaches. The information on potential epitopes obtained with the analysis

described here was also considered for the selection of suitable candidate molecules for comprehensive epitope analysis of Gly m 4 as described in the next section.

3.2.2 Rational mutagenesis and known epitope data

In addition to the peptide-based epitope analysis in the previous section a rational approach to select amino acids potentially involved in antibody binding based on the known three-dimensional structure of Gly m 4 was followed. Gly m 4 consists of 158 amino acids of which 79 amino acids are surface-exposed (with ≥20 $Å^2$ accessible surface area) comprising a total surface area of 6313.6 $Å^2$ from 9118.7 $Å^2$ of total protein surface. For subsequent mutational epitope analysis in which individual amino acids were analyzed for their impact in IgE binding the following criteria were considered to ensure maintenance of Gly m 4-type conformation: i) the surface-exposed area of individual amino acids must be larger than 111 $Å^2$ (Lins *et al.*, 2003). Thus alanines and glycines were not included. ii) proline was excluded from the analysis because of its high impact on secondary structure topology, iii) amino acid side chains must be capable of undergoing hydrogen bonds and/or electrostatic interactions (charged, uncharged, polar and aromatic amino acids), iv) side chains of iii) must not be involved in intramolecular interactions (H-bonds or salt bridges) and v) amino acids exposed to the intramolecular hydrophobic cleft were excluded.

In addition to the selection criteria above and results obtained by peptide-based epitope analysis (see 3.2.1), published data on single amino acids potentially involved in antibody binding of Bet v 1 or Bet v 1-homologs were evaluated and also considered (Scheurer *et al.*, 1999; Mirza *et al.*, 2000; Neudecker *et al.*, 2003a; Spangfort *et al.*, 2003; Holm *et al.*, 2004; Jahn-Schmid *et al.*, 2005; Wiche *et al.*, 2005; Wangorsch *et al.*, 2007; Gieras *et al.*, 2011; Hecker *et al.*, 2012). Taken together, a total number of 51 amino acids were considered suitable for a comprehensive epitope analysis of Gly m 4 (Table 11).

Table 11: Amino acids of Gly m 4 potentially involved in IgE binding.

	Amino acid		Amino acid		Amino acid
Non-polar	F3	**Polar**	T116	**Acidic**	D124
	F5		N128		E126
	Y156		Q129		D130
Polar	T4	**Acidic**	E6		E131
	N10		D7		D140
	T17		D25		E147
	T24		D27		D155
	N28		D35	**Basic**	K20
	S36		E44		K32
	S39		E59		K38
	N42		D60		K54
	N46		E71		K64
	S81		E75		K96
	S83		D92		K118
	T93		E95		K133
	T98		D100		K136
	N108		E120		K144

51 amino acids as constituents of putative epitopes of Gly m 4 for IgE were chosen according to epitope mapping, rational considerations and published data. Amino acids are sorted by the polarity and charge of their side chains.

In Figure 4, 51 amino acids selected for epitope analysis are shown. The candidate residues are distributed over the complete Gly m 4 primary structure (Figure 4 A). In total, 61% (31/51) of residues considered for putative IgE binding are charged (aspartate (11x), glutamate (10x) and lysine (10x)), 33% (17/51) are polar (threonine (6x), asparagine (6x), serine (4x) as well as glutamine (1x)) and 6% (3/51) show non-polar side chains (phenylalanine (2x) and tyrosine (1x)). The distribution of the selected candidate residues on the molecular surface of Gly m 4 is shown in Figure 4 B.

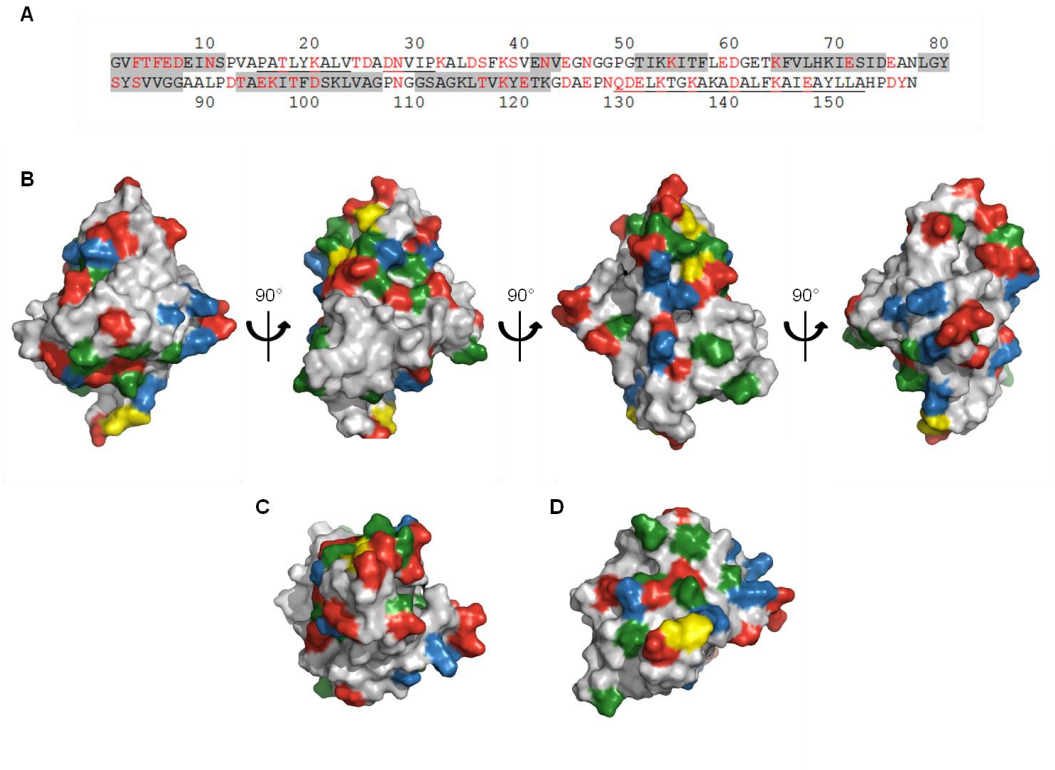


Figure 4: Amino acids of Gly m 4 selected for mutational epitope analysis.

(A) Gly m 4 amino acid sequence with 51 candidate residues colored in red. Secondary structures of β-strands (grey) and α-helices (underlined) are shown. (B) Surface representation of Gly m 4 with views turned in 90° rotation clockwise. (C) Top view. (D) Bottom view. Colored amino acids represent chosen 51 candidate residues: non-polar - yellow; polar - green; acidic - red; basic - blue. Pictures were generated with PyMOL surface presentation using Gly m 4 pdb file 2K7H.

3.3 Generation of a Gly m 4-specific epitope library for IgE

To analyze the potential impact of the 51 amino acids of Gly m 4 selected in the previous section on binding serum IgE of birch-associated soy allergic subjects, a recombinant model protein with Gly m 4-type conformation but no/low IgE-binding capacity was used. Norcoclaurine synthase (NCS) from the meadow rue (*Thalictrum flavum)* has Gly m 4-type protein conformation, low amino acid sequence homology to Gly m 4 (20.4% sequence identity) and no known allergenic properties (Ilari *et al.*, 2009; Berkner *et al.*, 2014). These characteristics make NCS an ideal model protein to analyze the impact of individual amino acids of Gly m 4 on serum IgE-binding. For this, single amino acids of NCS were replaced with structurally homologous amino acids of Gly m 4. Provided that the Gly m 4-type protein fold of NCS is not altered upon amino acid substitution, serum IgE-binding of such a NCS protein variant would directly point to a specific role of the substituted Gly m 4-specific amino acid in IgE interaction of Gly m 4 (Figure 5).

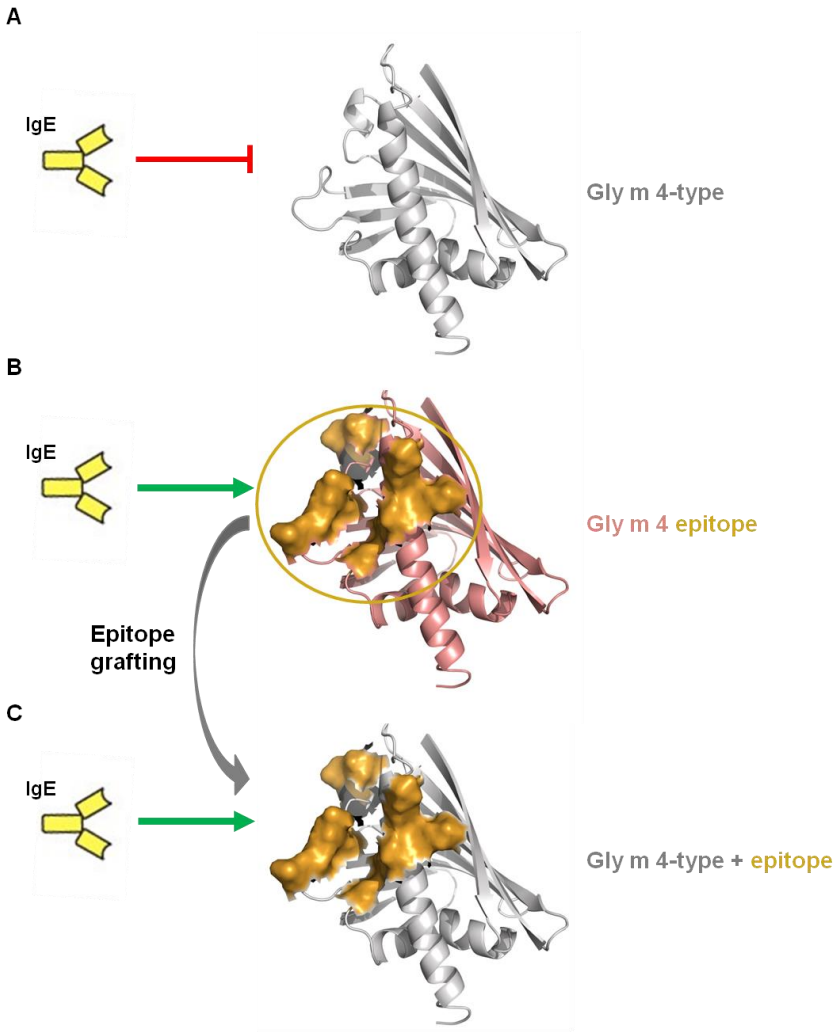


Figure 5: Principle of epitope grafting using NCS as Gly m 4-type model protein.

(A) Gly m 4-type model protein NCS (grey) with no/low IgE-binding capacity. (B) IgE antibodies bind to a specific epitope (orange) of Gly m 4 (red). (C) Transfer of amino acids from an IgE epitope of Gly m 4 to NCS induces IgE-binding of the Gly m 4-type NCS variant and enables the analysis of the impact of individual amino acids of Gly m 4 in serum IgE-binding. Pictures were generated with PyMol.

Two NCS model proteins were used in this study. Δ51NCS$_{N42/P49}$ was generated with a total length of 160 amino acids harboring two specific amino acid insertions/substitutions typical for Gly m 4-homologous allergens. The protein was also truncated N- and C-terminally to adjust it to a typical length of Bet v 1-type allergens (Nürnberg, 2014, Master thesis). Furthermore Δ51NCS$_{N42/P49}$ shows two gaps in C-terminal α-helix most probably due to two proline residues at positions 134 and 140 (pdb: 2VNE) (Ilari *et al.*, 2009). Therefore Δ51NCS$_{N42/P49/P134E/P140S}$ was generated in this study to replace the two prolines that do prevent the formation of a continuous C-terminal helix typical for Bet v 1-type proteins. The structural alignment of Δ51NCS$_{N42/P49}$, Bet v 1a and Gly m 4 is shown in Figure 6. With the two NCS

model proteins a library of recombinant NCS variants harboring potential epitope residues of Gly m 4 was generated.

Figure 6: Secondary structure and structural alignment of Δ51NCS$_{N42/P49}$, Bet v 1a and Gly m 4, respectively.

(A) Overlay of secondary structure topologies of Δ51NCS$_{N42/P49}$ (blue), Bet v 1a (green, pdb: 1BV1) and Gly m 4 (red, pdb: 2K7H). (B) Structural sequence alignment of Δ51NCS$_{N42/P49}$, Bet v 1a and Gly m 4 with β-strands (grey) and α-helices (underlined) indicated. Amino acid identities to Δ51NCS$_{N42/P49}$ are 19.5% (Bet v 1a) and 20.4% (Gly m 4), respectively. Overlay of Δ51NCS$_{N42/P49}$, Bet v 1a and Gly m 4 was generated with PyMol while EMBOSS Needle Pairwise Alignment tool provided by European Bioinformatics Institute was used for alignment of amino acid sequences.

A basic concept in this work is the understanding of a functional epitope as described by Dall'Acqua *et al.* where a single residue within a structural IgE epitope dominates the energetics of allergen-IgE binding (Dall'Acqua *et al.*, 1998). Therefore the grafting of a single amino acid from Gly m 4 onto non-IgE binding NCS model protein might mimic an antibody-binding site of Gly m 4 resulting in specific binding of serum IgE. In the rest of the study the procedure of grafting amino acid(s) which evoke IgE binding is considered as the generation of a (putative) functional epitope bearing in mind that a single residue *per se* does not bind IgE, but its introduction might adapt the molecular surface of NCS more similar to Gly m 4 thus enabling the interaction of such a modified NCS variant with Gly m 4-specific IgE antibodies (Figure 7).

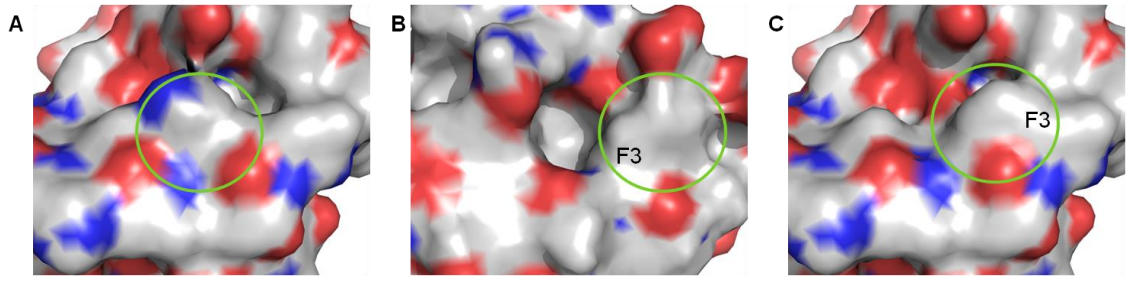


Figure 7: Structural changes upon grafting of amino acid F3 from Gly m 4 to NCS.

Surface representation of $\Delta 51NCS_{N42/P49}$ (A), Gly m 4 (B) and $\Delta 51NCS_{F3/N42/P49}$ (C). Green circle highlights grafted amino acid F3 in Gly m 4 and $\Delta 51NCS_{F3/N42/P49}$ as well as corresponding residue in $\Delta 51NCS_{N42/P49}$. Colored atoms represent: oxygen – red, nitrogen – blue. Pictures were generated with PyMOL surface presentation.

The structural-homologous grafting of the single amino acid F3 from Gly m 4 into NCS model protein not only changes the molecular surface at this position in NCS but also introduces structural changes in the arrangement of adjacent residues thereby leading to a potential epitope-like molecular surface that might result in serum IgE binding.

In this study, site-directed mutagenesis of NCS resulted in a total number of 48 different variants containing all 51 residues chosen as potential amino acids involved in binding of IgE (see Table 11). Most variants of the NCS library were generated with single amino acid substitutions while eight NCS variants comprised two substitutions each. Furthermore, one variant with five substituted residues ($\Delta 51NCS_{N42/E44/N46/P49/K54}$) was included in screening as it contains 75% of the experimentally determined epitope of mouse Bet v 1-specific monoclonal antibody BV16. This particular epitope of Bet v 1 is highly similar to the corresponding region of Gly m 4 and has been recently analyzed as corresponding Δ29NCS variant with Bet v 1-type protein fold as determined by CD and NMR spectroscopy and the capability to bind IgE of birch-allergic patients (Berkner *et al.*, 2014). All amino acids of NCS that are structurally homologous to the individual amino acids of Gly m 4 potentially involved in IgE-binding are listed in Table 12.

Table 12: Candidate residues of Gly m 4 and their corresponding amino acids of Δ51NCS.

	Amino acid			Amino acid	
	Gly m 4	**Δ51NCS**		**Gly m 4**	**Δ51NCS**
Non-polar	F3	K4	**Acidic**	E75	N78
	F5	I6		E120	H123
Polar	T4	V5		D124	E127
	N10	E11		D124/E126	E127/V129
	T17	D18		D130	P134
	T24	P26		E131	L135
	S36	A39		D140	A142
	S39	K42		E147	S149
	N46	D49	**Basic**	K20	T21
	S81	K84		K32	D34
	S83	Q86		K38	E41
	T98	M101		K54	D57
	N108	K111		K64	H67
	T116	S119		K118	E121
	N128	V132		K133	I136
	Q129	E133		K136	T138
Acidic	E6	H7		K144	D146
	D7	H8	**Mixed**	F5/D100	I6/T103
	D25	G27		N28/K32	K30/D34
	D25/D27	G27/A29		E44/N46	I47/D49
	D35	G38		D92/T93	L95/G96
	E59	P62		E95/K96	T98/Y99
	D60	G63		D155/Y156	S157/S159
	E71	I74		N42/E44/N46/P49/K54	5x

Amino acids are sorted by their specific side chain characteristics as shown in Table 11. Variants with two amino acid substitutions belonging to different groups of amino acids are listed as mixed residues. Each aa of Gly m 4 as well as the corresponding structural homologous residue in Δ51NCS that was replaced for the Gly m 4-specific residue is shown. In total 48 Δ51NCS variants were generated by site-specific mutagenesis with $\Delta51NCS_{N42/P49}$ or $\Delta51NCS_{N42/P49/P134E/P140S}$ as template.

3.4 Screening of the NCS protein library for serum IgE-binding

3.4.1 A murine monoclonal antibody-binding epitope of Bet v 1 also binds serum IgE

Gly m 4-specific amino acids N42, E44, N46, P49 and K54 grafted onto NCS model protein were shown to induce binding of IgE of sera from birch pollen-allergic subjects to the Δ29NCS variant (Berkner *et al.*, 2014). Furthermore IgE cross-reactivity between Bet v 1-homologous food allergens from soy and hazelnut harboring these amino acids was observed. To analyze sera of birch allergic patients with (group 1) and without (group 2) soybean allergy for IgE-binding to the epitope of

monoclonal antibody BV16, the optimized recombinant variant $\Delta51NCS_{N42/E44/N46/P49/K54}$ was used. Observed binding of IgE antibodies compared to model protein $\Delta51NCS_{N42/P49}$ indicated presence of serum IgE able to bind grafted residues. Immunoblots with recombinant $\Delta51NCS_{N42/P49}$ and $\Delta51NCS_{N42/E44/N46/P49/K54}$ variants and sera of group 1 and 2 are shown in Figure 8.

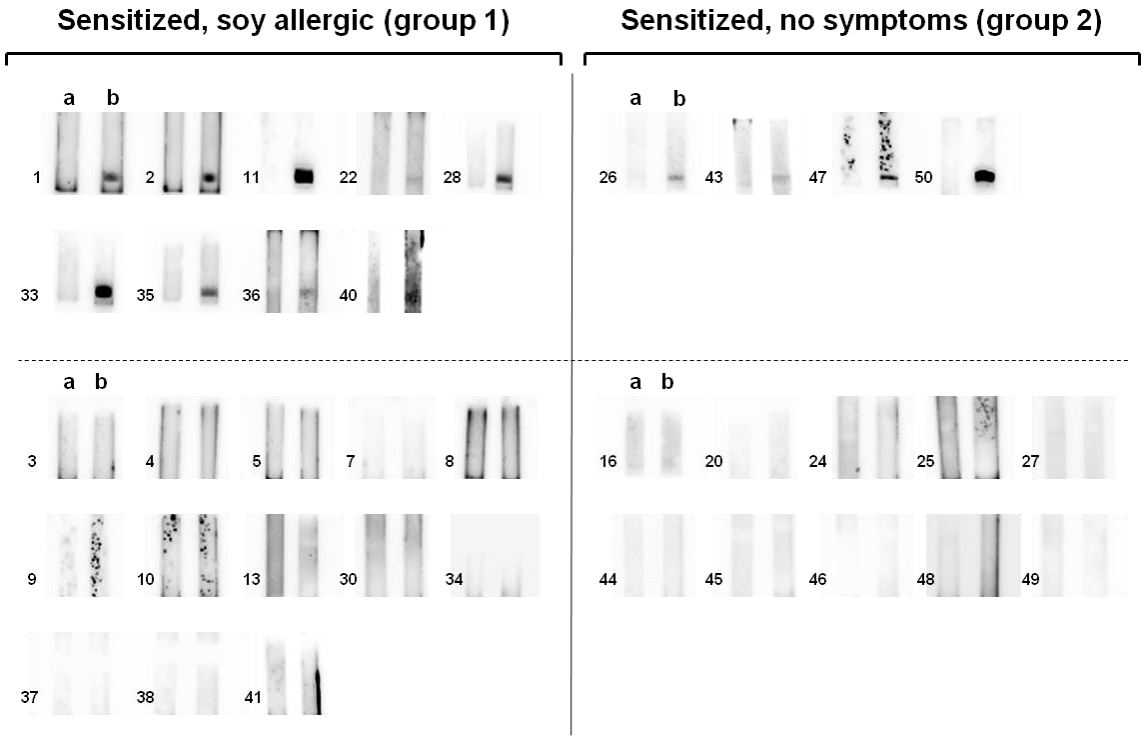


Figure 8: $\Delta51NCS_{N42/E44/N46/P49/K54}$ binds IgE of several sera of birch allergic patients with and without soybean allergy.

Patients sensitized to Bet v 1 and Gly m 4, with (group 1) and without (group 2) clinically confirmed soybean allergy were analyzed in immunoblot. IgE binding to reference model protein $\Delta51NCS_{N42/P49}$ (lane a) and epitope-presenting variant $\Delta51NCS_{N42/E44/N46/P49/K54}$ (lane b) is shown. Each serum was used in a 1:20 dilution with patient number indicated at each blot. Patients are sorted for positive IgE response (upper panel) and no IgE response (lower panel).

Several patients had IgE antibodies against the particular epitope presented by the $\Delta51NCS_{N42/E44/N46/P49/K54}$ variant. In total, 41% (9/22) of patients sensitized and allergic to soy (group 1) and 29% (4/14) of patients sensitized but without soybean allergy (group 2) showed binding of serum IgE to $\Delta51NCS_{N42/E44/N46/P49/K54}$ variant, respectively. In comparison to $\Delta51NCS_{N42/P49}$ reference, $\Delta51NCS_{N42/E44/N46/P49/K54}$ harbors three additional grafted residues of Gly m 4, namely E44, N46 and K54 which seem to be important for IgE binding at least in some of the patients' sera analyzed. Frequency of IgE binding to $\Delta51NCS_{N42/E44/N46/P49/K54}$ combining both groups is 36% (13/36) in immunoblot, which is much lower compared to previous

ImmunoCAP™ analyses showing 71% (50/70) of birch pollen-allergic subjects with IgE binding to grafted BV16-binding epitope (Berkner *et al.*, 2014). One possible explanation might be the higher sensitivity of ImmunoCAP™ compared to immunoblot analyses.

3.4.2 Screening of a NCS protein library for functional IgE epitopes

To identify further functional IgE epitopes of Gly m 4, the generated library of NCS proteins was used (see 3.3). Thus each of the 48 recombinant NCS variants was expressed in small-scale culture and bacterial lysates were analyzed for IgE binding with sera of soy allergic subjects (group 1) and sera of subjects non-allergic to soy (group 2) in immunoblots. As IgE epitopes of Bet v 1 and its homologues are conformational, observed IgE binding to NCS variants under these conditions would point to stably Gly m 4-type refolding of NCS variants. Figure 9 shows an exemplary SDS-PAGE of several NCS variants after cell lysis with crude bacterial lysates plotted which were also used in immunoblot analysis below.

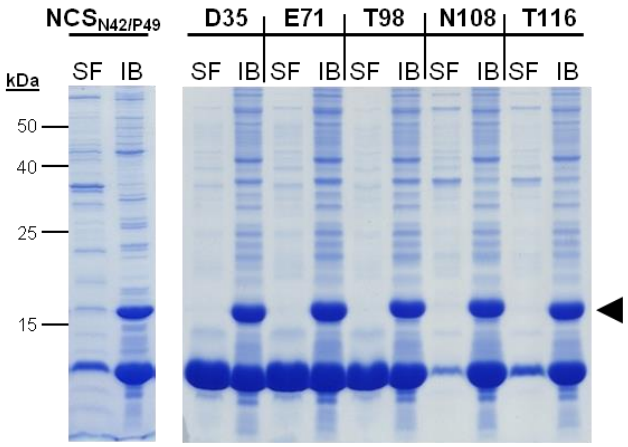


Figure 9: Expression of reference model protein Δ51NCS$_{N42/P49}$ and Δ51NCS$_{N42/P49}$ variants in *E. coli*.

Coomassie-stained gel of bacterial lysates of *E. coli* expressing recombinant model protein Δ51NCS$_{N42/P49}$ and Δ51NCS$_{N42/P49}$ variants presenting Gly m 4-specific amino acids D35, E71, T98, N108 and T116 upon SDS-PAGE. Shown are soluble fractions (SF) and inclusion bodies (IB) of small-scale expressions in *E. coli* obtained by cell lysis using BugBuster extraction kit. Arrow heads (◄) mark expressed target proteins.

All recombinant Δ51NCS$_{N42/P49}$ variants were expressed almost exclusively as non-soluble proteins in inclusion bodies. Therefore, in immunoblotting estimated comparable amounts of crude inclusion body protein fraction of each NCS variant was analyzed. As reference protein, bacterial lysate of Δ51NCS$_{N42/P49}$ showing no or

very low IgE-binding capacity was used. By quantitative comparison of IgE signal intensities observed with the NCS variants to reference protein $\Delta51NCS_{N42/P49}$, amino acids potentially involved in Gly m 4-specific IgE binding could be identified. With this initial screening a patient-specific profile of potential Gly m 4-specific functional IgE epitopes could be identified which can be used as initial results for detailed analyses afterwards. As an example, Figure 10 shows the initial screening of all NCS variants together with serum 33 (group 1) and 49 (group 2).

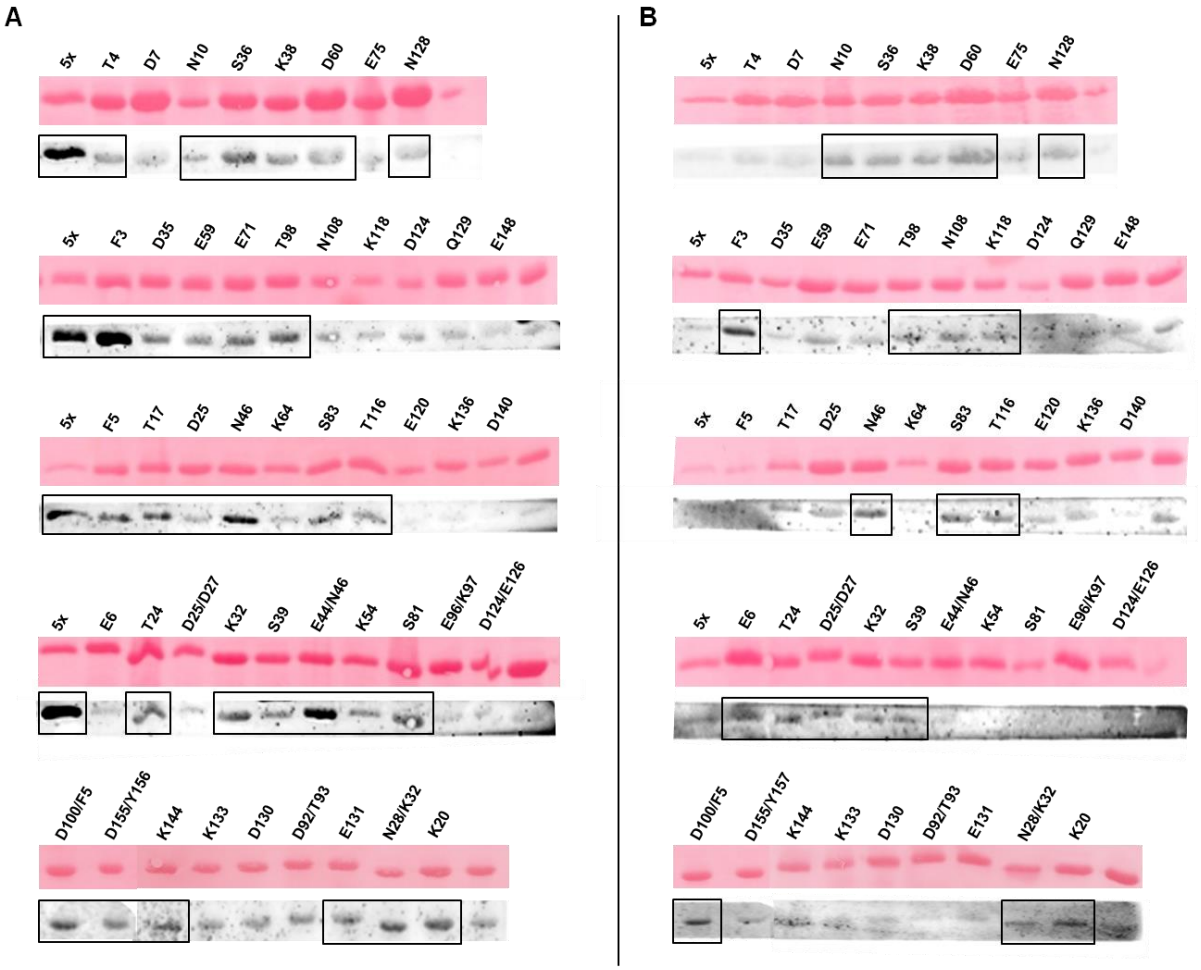


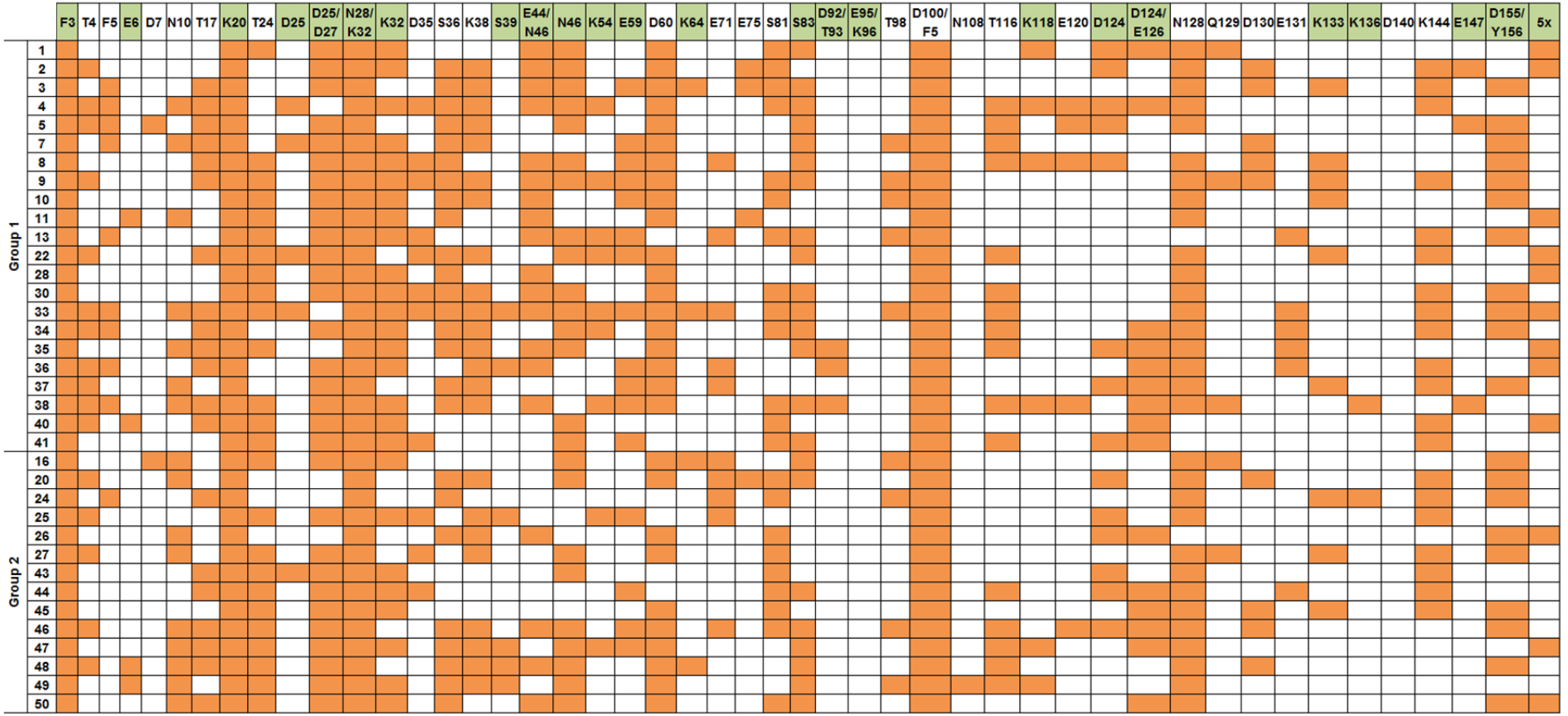
Figure 10: Screening of NCS library with patient's serum No. 33 (A) and 49 (B).

Shown are Ponceau S-stained $\Delta51NCS_{N42/P49}$ proteins as well as corresponding IgE signals obtained by immunoblotting. Sera of patient No. 33 (A, sensitized and soy allergic, group 1) and 49 (B, sensitized, no soybean allergy, group 2) were used in a 1:10 dilution. Gly m 4-specific amino acid(s) presented by $\Delta51NCS_{N42/P49}$ variants are indicated while 5x resembles $\Delta51NCS_{N42/E44/N46/P49/K54}$ variant. Protein used at the very right end of each blot resembles $\Delta51NCS_{N42/P49}$ control protein necessary for analysis of each blot. NCS variants with positive IgE binding signals compared to control protein are highlighted.

For patient 33 a total number of 31 putative IgE-binding NCS variants could be identified resulting in 31 putative functional IgE epitopes. Furthermore, serum 33 has IgE antibodies directed against $\Delta51NCS_{N42/E44/N46/P49/K54}$ variant, confirming results

obtained by immunoblotting with purified protein (cf. Figure 8). In contrast no sIgE directed against $\Delta 51NCS_{N42/E44/N46/P49/K54}$ could be detected in serum of patient 49 but a total number of 20 putative functional IgE epitopes. In this way all 36 sera of group 1 and 2 were screened and the results of all identified putative functional IgE epitopes are summarized in Table 13.

Table 13: Putative functional IgE epitopes of Gly m 4.



48 Δ51NCS$_{N42/P49}$ variants presenting indicated Gly m 4-specific amino acid(s) were analyzed in immunoblotting with sera of birch allergic patients with (22 sera, group 1) or without (14 sera, group 2) soybean allergy. IgE signals of variants were compared with Δ51NCS$_{N42/P49}$ control protein on each blot and variants with positive IgE binding signal are colored in orange. Gly m 4-specific amino acids presented by Δ51NCS$_{N42/P49}$ model protein are listed from N- (left) to C-terminus (right) with amino acids identical in Gly m 4 and Bet v 1a colored in green. 5x represents Δ51NCS$_{N42/E44/N46/P49/K54}$ variant. Patients are listed in group 1 and group 2 according to clinical reactivity to soybean shown in Table 9.

In total 48 NCS variants were screened with 36 sera comprising group 1 and 2. 96% (46/48) of NCS variants bound IgE from sera of at least one patient compared to reference protein $\Delta51NCS_{N42/P49}$ resulting in the identification of 46 putative functional IgE epitopes. 48% (22/46) of the NCS variants generated presented amino acids homologous between Gly m 4 and Bet v 1a, while 52% (24/46) of identified putative functional IgE epitopes were Gly m 4-specific. Only $\Delta51NCS_{N42/P49/E95/K96}$ and $\Delta51NCS_{N42/P49/D140}$ did not bind IgE from any of the sera analyzed. Average number of putative functional epitopes for each serum is comparable between both groups showing 21 and 18 identified putative epitopes for patients with and without clinical soy allergy, respectively. The range of 12 to 32 putative functional epitopes identified for patients with soybean allergy was larger than 14-26 putative functional epitopes for patients sensitized but not allergic to soy. Two NCS variants $\Delta51NCS_{F5/N42/P49/D100}$ and $\Delta51NCS_{F3/N42/P49}$ bound serum IgE of all sera tested, whereas NCS variants $\Delta51NCS_{K20/N42/P49}$, $\Delta51NCS_{D25/D27/N42/P49}$, $\Delta51NCS_{N28/K32/N42/P49}$, $\Delta51NCS_{N42/P49/D60}$ and $\Delta51NCS_{N42/P49/N128}$ bound serum IgE of at least 75% of all sera. 39% (18/46) of the NCS library bound serum IgE of more than 50% of the study population. Reference protein $\Delta51NCS_{N42/P49}$ showed weak IgE binding with sera of several patients most likely because of identical amino acids between $\Delta51NCS_{N42/P49}$ and Gly m 4 or Bet v 1a primary sequence with possible importance for interaction with IgE antibodies. In total $\Delta51NCS_{N42/P49}$ shares 31 and 35 residues with Gly m 4 and Bet v 1a respectively, whereas 15 and 16 amino acids might be involved in IgE binding because of their side chains (E, F, H, K, N, S, T, Y). A complete analysis of the IgE binding frequency of study groups 1 and 2 is shown in Figure 11 A and B.

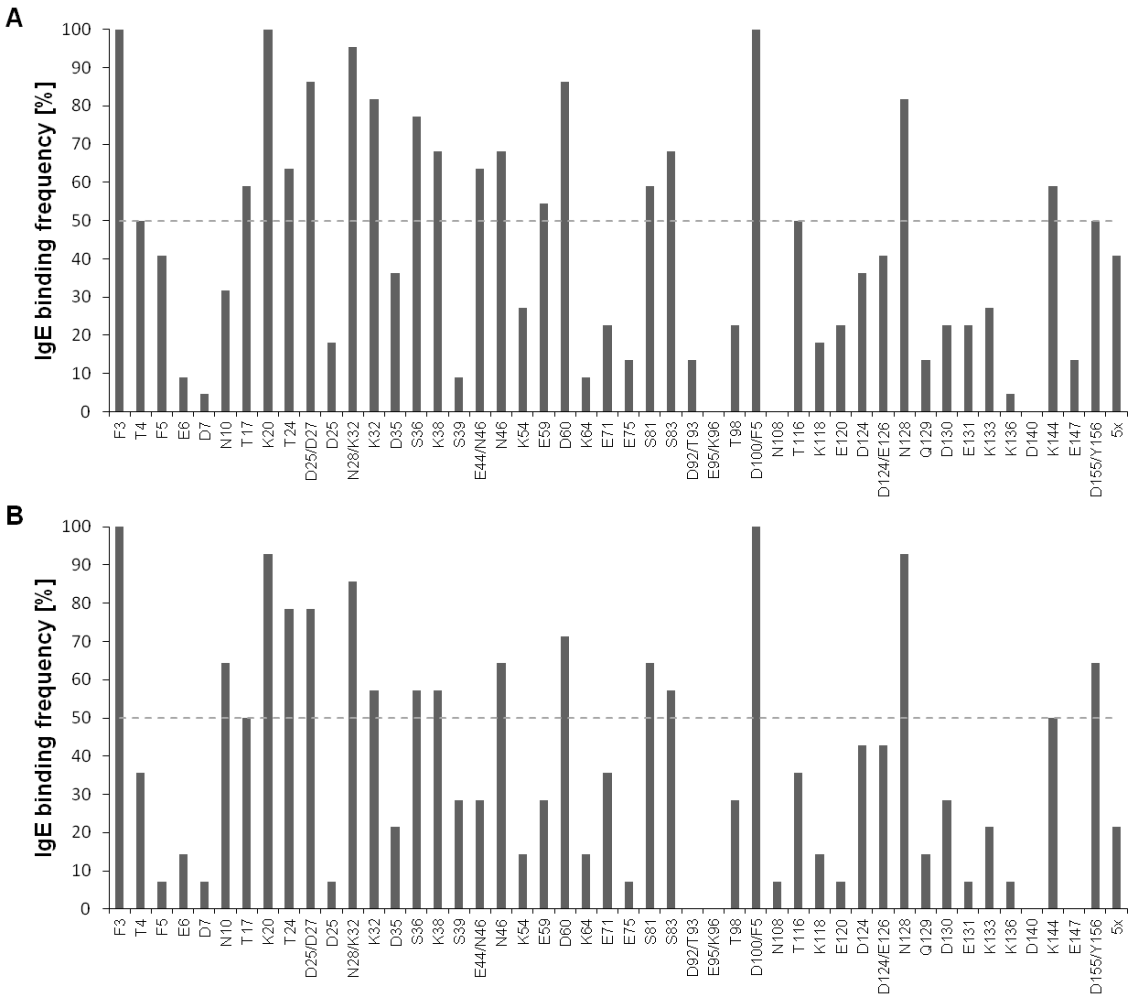


Figure 11: IgE binding frequencies of Δ51NCS$_{N42/P49}$ variants presenting Gly m 4-specific amino acid(s) are highly individual.

IgE binding frequency in % of all tested Δ51NCS$_{N42/P49}$ variants presenting Gly m 4-specific amino acid(s) with subjects sensitized and soy allergic (A) and subjects without soybean allergy (B). Variants are listed from N- (left) to C-terminus (right) with respective Gly m 4 residues grafted. 5x represents Δ51NCS$_{N42/E44/N46/P49/K54}$ variant. Dotted line represents an IgE binding frequency of 50%.

Two variants were identified as positive only with sera of one group of patients. NCS variants Δ51NCS$_{N42/P49/D92/T93}$ and Δ51NCS$_{N42/P49/E147}$ bound IgE only with three sera of patients sensitized and soy allergic each. In contrast Δ51NCS$_{N42/P49/N108}$ bound only IgE antibodies in one serum of a patient without soybean allergy. Despite of these three NCS variants there were no obvious different characteristics observable regarding the distribution of Gly m 4-specific putative functional IgE epitopes between birch-soy allergic and Gly m 4-sensitized patients only. Rather, every patient exhibited an individual pattern of putative functional IgE epitopes which did not correlate to specific clinical symptoms. In Figure 12 all identified putative functional epitopes, as well as functional epitopes identified with patients 33 and 49 are shown on Gly m 4 protein.

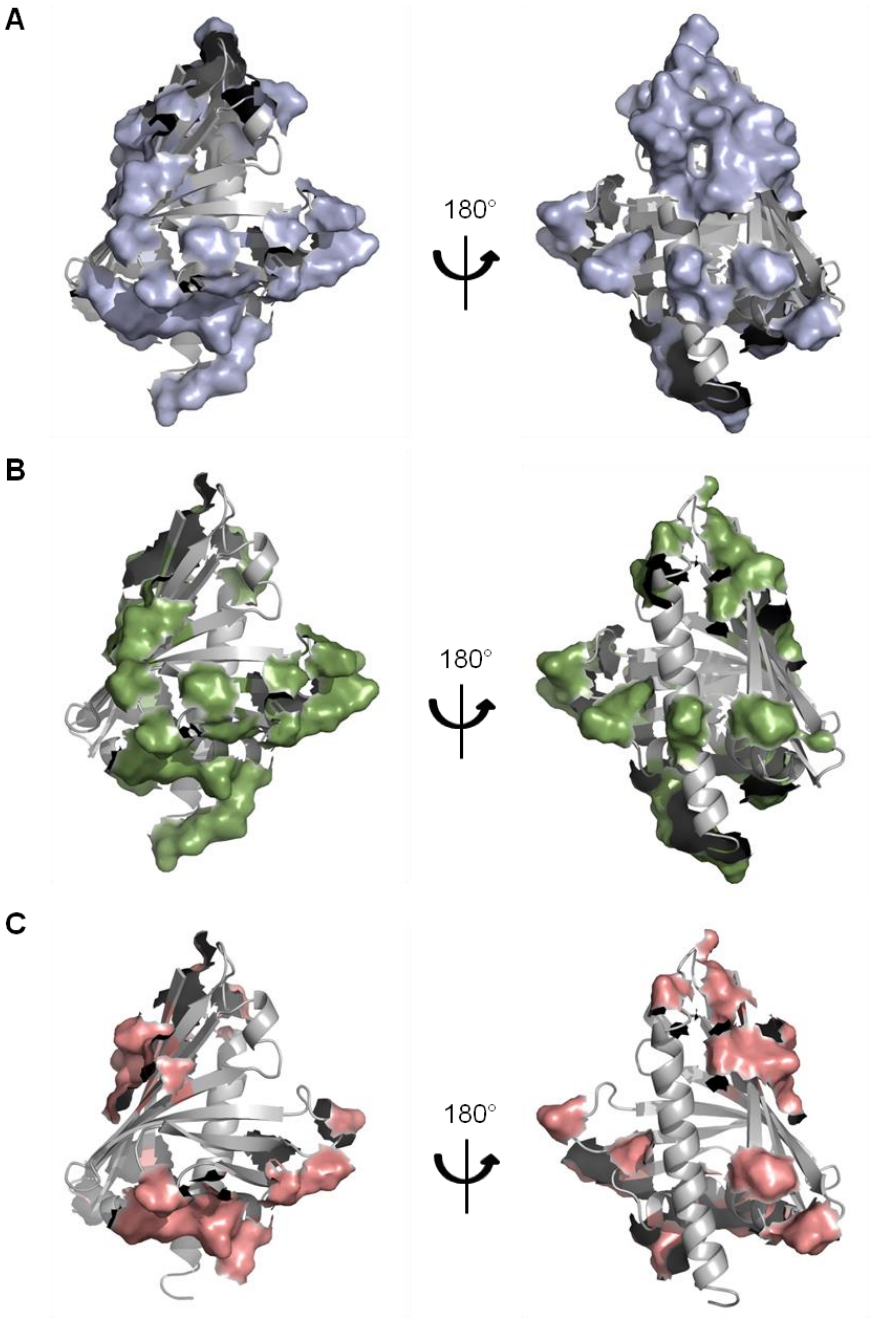


Figure 12: Putative functional IgE epitopes of Gly m 4.

Surface areas of putative functional IgE epitopes of Gly m 4 identified by screening of the Δ51NCS$_{N42/P49}$ epitope library with representative sera of patients with (patient 33) and without (patient 49) allergy to soy. In total 47 putative functional epitopes were found (A). 31 putative functional IgE epitopes for patient 33 (B) and 20 putative functional IgE epitopes for patient 49 (C) are indicated. Protein models were generated with PyMOL surface presentation using Gly m 4 pdb file 2K7H.

Putative functional IgE epitopes (see definition in section 3.3 and Figure 7) identified by screening the NCS library were distributed all over Gly m 4 molecule and covered about 45% (4112.8 Å^2) of surface accessible protein surface area. Only two larger patches not involved in IgE binding in this analysis were presented by surface areas consisting of parts of β-strands 4 and 5 as well as β-strands 6 and 7. These

segments of the Gly m 4 sequence showed many glycine, alanine and valine residues which putatively do not interact with antibodies, as well as amino acids K69 or K114 homologous in NCS model protein and Gly m 4. When comparing NCS binding of sera of patients 33 and 49, respectively, differences as well as similarities in the distribution of putative functional IgE epitopes became evident. For patient 33, 31 putative functional epitopes covering 26% (2342.6 $Å^2$) of total protein surface were found. For patient 49, 20 putative functional epitopes were identified resulting in coverage of 17% (1528.2 $Å^2$) of the Gly m 4 surface. This causes larger surface areas not involved in IgE binding for patient 49 compared to patient 33. Concordant between both patients are surface patches around short α-helices or β-sheets 1, 6 and 7. In detail these are amino acids F3, N10 and K20 for short α-helices and N28/K32, T98, D100/F5 and T116 for β-sheets. The initial results obtained by screening of NCS with immunoblot are further analyzed in subsequent sections of this study.

All identified amino acids were clustered to six putative IgE-binding areas since such spatial arrangement results in a maximum number of virtual epitopes on total Gly m 4 surface enabling a simultaneous binding of antibodies to at least two areas (Figure 13) which is a prerequisite for the allergic reactions initiated via Gly m 4-induced IgE cross-linking on activated basophils and mast cells (cf. Figure 1). Residues at the intersection of two virtual epitope areas might represent amino acids belonging to more than one of the six hypothetical IgE-binding areas. These putative functional IgE epitopes (I-VI) cover 3-11% of total protein surface and span a molecular surface of 311.4-967.3 $Å^2$ each while consisting of 3-10 amino acids. Epitope I consists of amino acids F3, T4, F5, E6, D7, D124, E126, N128 and Q129 comprising portions of Gly m 4 N-terminus and loop 9 connecting last β-strand with long C-terminal α-helix. In close proximity epitope II is located with most residues belonging to C-terminal helix (D130, E131, K133, K136, K144 and E147) expanded with two asparagines N10 and N108. Epitope III consists of amino acids T17, K20, T24, D25, D27, N28 and K32 forming main parts of both short α-helices as well as E75 and residues D155 and Y156 at C-terminus of Gly m 4 sequence. Residues D35, S36, K38, S39, E59, D60, K64, D92 and T93 together comprise epitope IV mainly located at loops 3 and 5. Epitope V presents residues grafted onto $\Delta 51NCS_{N42/E44/N46/P49/K54}$ variant showing parts of structurally solved BV16 epitope located in the conserved p-loop region. Formed by parts of β-sheets 4-7 epitope VI constitutes amino acids E71, S81

S83, T98, D100, T116, K118 and E120. Secondary structure of Gly m 4 with putative IgE epitopes I-VI is shown in Figure 13.

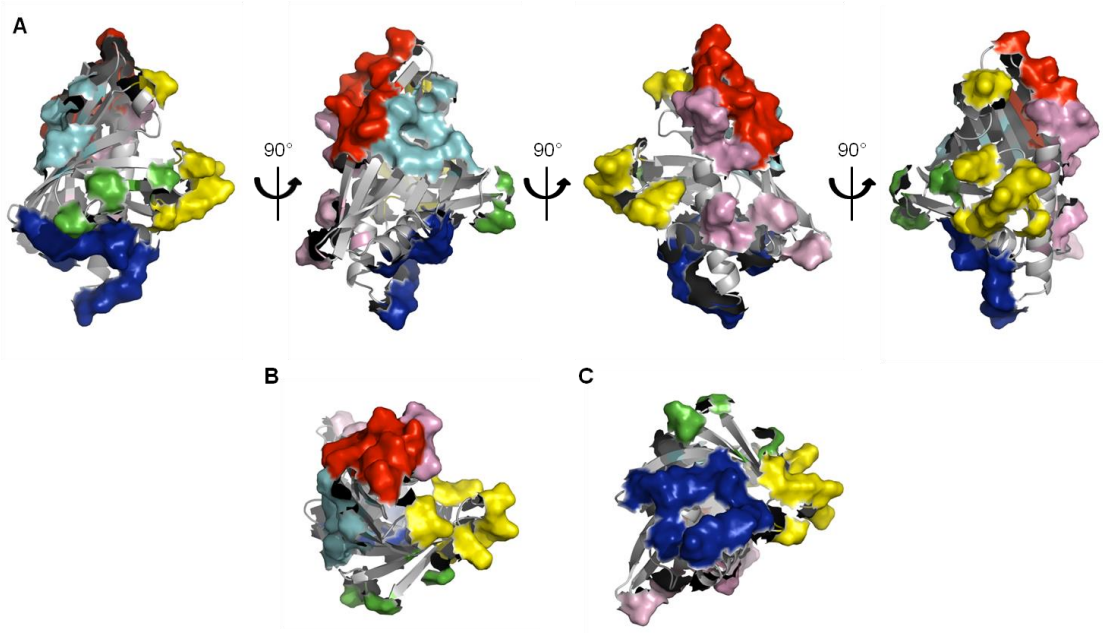


Figure 13: Six putative IgE epitopes of Gly m 4.

Putative functional IgE epitopes of Gly m 4 were combined. (A) Surface representation of six (I-VI) identified putative IgE epitopes of Gly m 4 with views turned in 90° rotation clockwise. (B) Top view. (C) Bottom view. Colors represent putative IgE epitopes: I (red) – F3, T4, F5, E6, D7, D124, E126, N128, Q129; II (dark pink) – N10, N108, D130, E131, K136, K144, E147; III (blue) – T16, K20, T24, D25, D27, N28, K32, E75, D155, Y156; IV (yellow) – D35, S36, K38, S39, E59, D60, K64, D92, T93; V (green) –E44, N46, K54; VI (cyan) – E71, S81, S83, T98, D100, T116, K118, E120. Pictures were generated with PyMOL surface and cartoon presentation using Gly m 4 pdb file 2K7H.

IgE binding frequencies of the six putative epitopes calculated by individual NCS variants recognized by serum IgE most frequently was 100% (36/36) for epitopes I, III, and VI, 81% (29/36) for epitope IV, 67% (24/36) for epitope V and 56% (20/36) for epitope II. A summary of putative epitopes with assignment of all identified functional IgE epitopes as well as possible combinations of six epitopes allowing simultaneous binding of two IgE antibodies is shown in Table 14 A and B.

Table 14: Putative functional IgE epitopes of Gly m 4 and their combinations allowing simultaneous binding of IgE antibodies.

A

IgE epitope	Amino acids	Number [aa]	Frequency [%]
I	F3, T4, F5, E6, D7, D124, E126, N128, Q129	9	100
II	N10, N108, D130, E131, K133, K136, K144, E147	8	56
III	T17, K20, T24, D25, D27, N28, K32, E75, D155, Y156	10	100
IV	D35, S36, K38, S39, E59, D60, K64, D92, T93	9	81
V	E44, N46, K54	3	67
VI	E71, S81, S83, T98, D100, T116, K118, E120	8	100

B

IgE epitope	I	II	III	IV	V	VI
I	-					
II	-	-				
III	+	+	-			
IV	+	+	+	-		
V	+	+	-	+	-	
VI	-	+	+	+	+	-

(A) Putative epitopes I-VI with all amino acids listed. IgE binding frequencies of the six putative epitopes were calculated by individual NCS variants recognized by serum IgE most frequently (substituted amino acids are underlined). Colors correspond to epitope patches shown in Figure 13. (B) Possibility of two putative epitopes binding their respective IgE antibodies is indicated with (+) or (-) whereas (+) indicates that simultaneous binding of two IgE antibodies is possible and (-) indicates two epitopes which are in close proximity to each other preventing simultaneous binding of IgE antibodies.

3.5 NCS protein variants with grafted epitope patches

3.5.1 Generation and characterization of multiple NCS variants

To experimentally verify the putative functional IgE epitopes identified in the previous section a more thorough analysis of purified recombinant NCS variants was necessary. For this, NCS variants combining individual and multiple potential IgE binding sites were generated (Table 15). A Gly m 4-type secondary structure of the NCS variants was evaluated by circular dichroism (CD) analysis of the purified model

proteins. However only eleven (variants 1-11) out of 14 variants generated showed Gly m 4-type CD spectra and were included in the further epitope analyses.

Table 15: Gly m 4 epitopes presented by recombinant NCS variants presenting Gly m 4-specific amino acid(s).

ΔNCS$_{N42/P49}$ variants	Grafted amino acids	Epitope patches
1	F3, F5, E6, D124, E126	I
2	D25, D27, N28, V29, K32, A152, D155, Y156	III
3	K38, S39, E59, D60, K64	IV
4	E44, N46, K54	V
5	N10, D130, E131, K133, K136, K144, E147, A148	II
6	K32, E44, N46, K54, D60, E120, E126	I, III, IV, V, VI
7	F3, K38, D60, N128	I, IV
8	T24, S36, S81, D124	I, III, IV, VI
9	T4, T17, N46, S83	I, III, V, VI
10	F5, T17, T24, S36, K64, S81, K118, D124, K133	I, II, III, IV, VI
11	T4, T17, K32, N46, E75, S83, D100, D124, E126, K144	I, II, III, V, VI
12	F3, N10, N28, K32, K38, D60, N128	I, II, III, IV
13	D7, D25, D27, D35, E44, N46, K54, E71, E120, E147	I, II, III, IV, V, VI
14	F5, K20, K38, E59, D60, D100, Q129, D155, Y156	I, III, IV, VI

Δ51NCS$_{N42/P49}$ variants 1-5 and 7-14 as well as Δ29NCS$_{N42/P49}$ variant 6 with grafted amino acids from six putative Gly m 4 epitopes for IgE (color code as shown in Table 14).

In total, three to ten amino acids were grafted from Gly m 4 to NCS model protein to create NCS variants with putative IgE-binding epitopes. Five variants displayed amino acids corresponding to one individual IgE epitope patch (I, II, III, IV or V) identified previously (see Table 14 A). Remaining six NCS proteins displayed combinations of different IgE binding sites identified. These multiple epitope variants combine 2-5 different epitope patches, respectively. Except Δ29NCS$_{K32/E44/N46/K54/D60/E120/E126}$ which is based on Δ29NCS$_{N42/P49}$ template, all further NCS variants were generated using Δ51NCS$_{N42/P49}$ as model protein. Subsequently, all NCS variants are designated as ΔNCS$_{N42/P49}$_1 to ΔNCS $_{N42/P49}$_11

according to their description listed in Table 15. In Figure 14 secondary structures of Δ51NCS$_{N42/P49}$ and variants of model protein are shown with functional epitopes indicated.

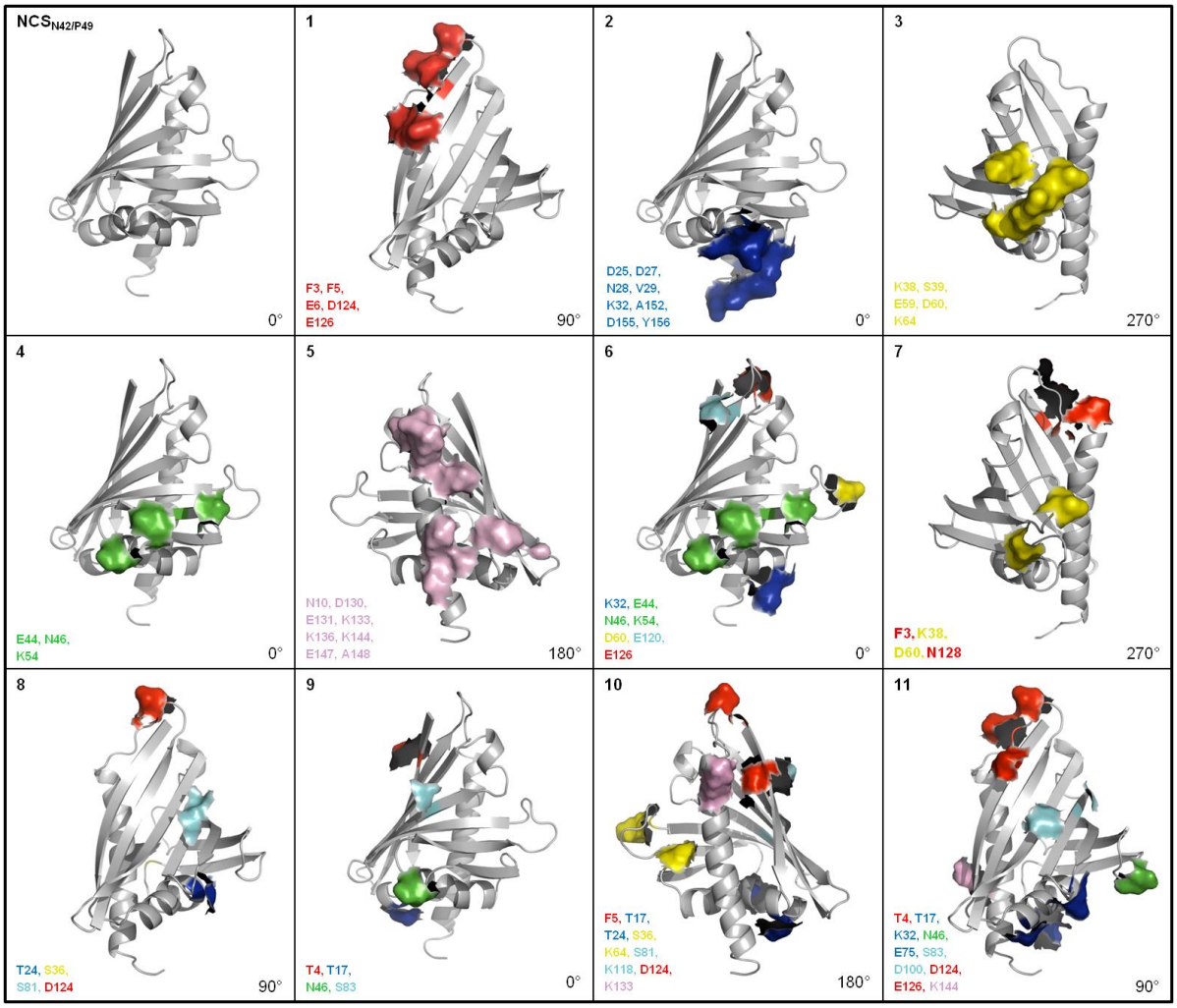


Figure 14: Secondary structure topologies of Δ51NCS$_{N42/P49}$ model proteins presenting individual and multiple IgE binding sites of Gly m 4.

ΔNCS$_{N42/P49}$ variants 1-5 present individual epitopes while variants 6-11 combine different IgE epitopes of Gly m 4. The amino acids are colored according to their epitope affiliation shown in Figure 13.

In total, all eleven putative IgE epitopes cover 35% (3199.4 Å^2) of total protein surface but each single epitope only matches a small surface patch of NCS (242.8-861.1 Å^2). For analysis of their IgE-binding potential all ΔNCS$_{N42/P49}$ variants were expressed and purified to homogeneity from *E. coli*. Purification was performed as described in 2.4.5 using affinity chromatography with C-terminal His-tag characteristic for pET29a vector followed by size exclusion chromatography. Total protein yield was about 20 mg out of 1 L of expression culture. Figure 15 shows

purified Δ51NCS$_{N42/P49}$ model protein as well as eleven ΔNCS$_{N42/P49}$ variants generated.

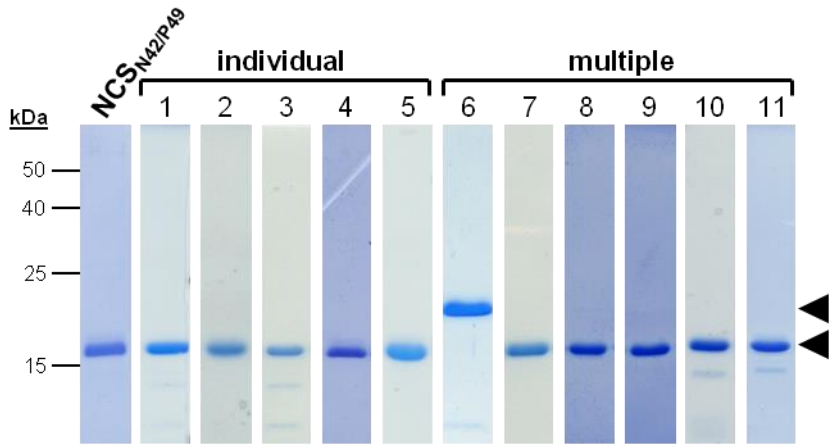


Figure 15: Purity of Δ51NCS$_{N42/P49}$ and ΔNCS$_{N42/P49}$ variants.

Purified ΔNCS$_{N42/P49}$ variants presenting individual or multiple IgE epitopes of Gly m 4 were analyzed in Coomassie-stained gels upon SDS-PAGE. Lanes 1 to 11 represent corresponding ΔNCS$_{N42/P49}$_1-11 variants. Arrow heads (◄) mark recombinant target proteins:
ΔNCS$_{N42/P49}$_1=Δ51NCS$_{F3/F5/E6N42/P49/D124/E126}$,
ΔNCS$_{N42/P49}$_2=Δ51NCS$_{D25/D27/N28/V29/K32/N42/P49/A152/D155/Y156}$,
ΔNCS$_{N42/P49}$_3=Δ51NCS$_{K38/S39/N42/P49/E59/D60/K64}$,
ΔNCS$_{N42/P49}$_4=Δ51NCS$_{N42/E44/N46/P49/K54}$,
ΔNCS$_{N42/P49}$_5=Δ51NCS$_{N10/N42/P49/D130/E131/K133/K136/K144/E147/A148}$,
ΔNCS$_{N42/P49}$_6=Δ29NCS$_{K32/N42/E44/N46/P49/K54/D60/E120/E126}$,
ΔNCS$_{N42/P49}$_7=Δ51NCS$_{F3/K38/N42/P49/D60/N128}$,
ΔNCS$_{N42/P49}$_8=Δ51NCS$_{T24/S36/N42/P49/S81/D124}$,
ΔNCS$_{N42/P49}$_9=Δ51NCS$_{T4/T17/N42/N46/P49/S83}$,
ΔNCS$_{N42/P49}$_10=Δ51NCS$_{F5/T17/T24/S36/N42/P49/K64/S81/K118/D124/K133}$,
ΔNCS$_{N42/P49}$_11=Δ51NCS$_{T4/T17/K32/N42/N46/P49/E75/S83\ /D100/D124/E126/K144}$.

Variants showed high purity with minor degradation products according to SDS-PAGE in five of purified ΔNCS$_{N42/P49}$ variants (ΔNCS$_{N42/P49}$_1, 3, 6, 9 and 10). PR-10-type secondary structures of the ΔNCS$_{N42/P49}$ variants were observed by CD spectroscopy (Figure 16).

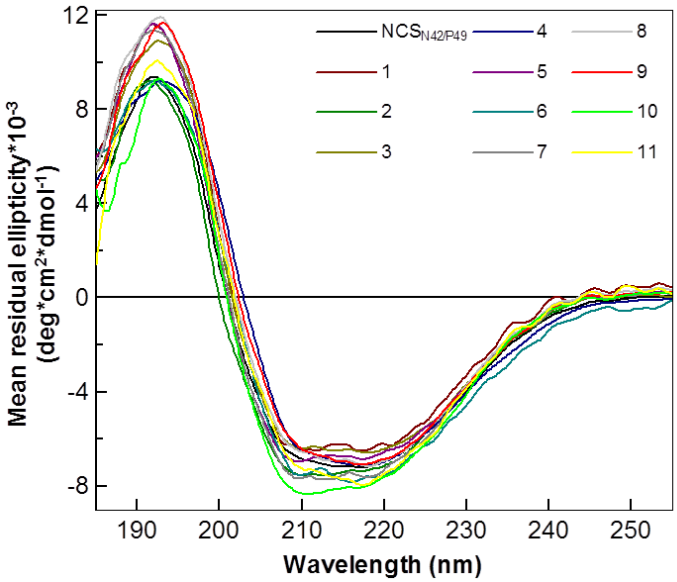


Figure 16: Circular dichroism of ΔNCS$_{N42/P49}$ variants.

CD spectra of Δ51NCS$_{N42/P49}$ and eleven corresponding ΔNCS$_{N42/P49}$ variants each recorded at 20 °C and 10 µM protein concentration in 10 mM potassium phosphate buffer at pH 7.4. Spectra are presented as mean residual ellipticity*10^{-3} at a given wavelength and are baseline corrected.

All recombinant ΔNCS$_{N42/P49}$ variants showed comparable secondary structure elements with Δ51NCS$_{N42/P49}$ model protein as determined with CD spectroscopy. Spectra indicate a high content of both α-helical regions and β-strands. Some minor differences between variants were observable especially around maxima at 210 and 194 nm. This might result from minor changes in protein concentration or small impurities left in protein samples. All in all, CD spectra of the ΔNCS$_{N42/P49}$ variants suggested that the amino acid substitutions did not induce relevant alterations to the Gly m 4-type secondary structure of the proteins.

3.5.2 Analysis of grafted epitope patches for binding serum IgE

All purified ΔNCS$_{N42/P49}$ variants with Gly m 4-type CD spectra were tested for serum IgE binding in a dot blot analysis. As reference proteins Δ51NCS$_{N42/P49}$ as well as Δ29NCS$_{N42/P49}$ (Berkner *et al.*, 2014) were used showing no or very low IgE binding to sera from sensitized subjects. Figure 17 A shows a sketch of the dot blot set-up with ΔNCS$_{N42/P49}$ variants analyzed for IgE binding with each patients' serum. Figure 17 B shows nine examples of dot blots obtained by incubation with sera of patients 11, 30, 33, 36, 37, and 41 for subjects sensitized and soy allergic (group 1)

as well as patients 20, 26 and 49 for subjects with no related soybean allergy (group 2), respectively, as well as non-atopic serum control.

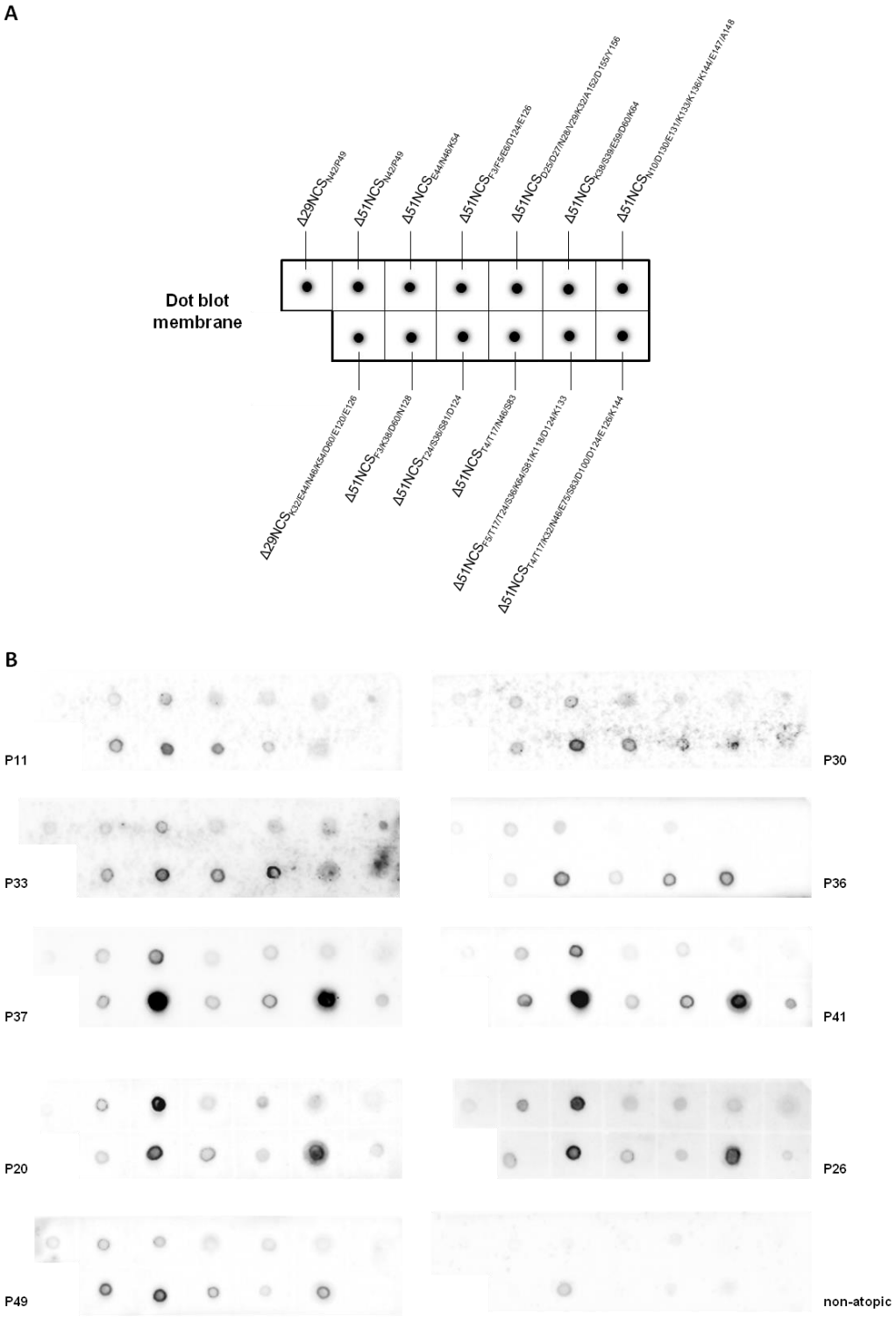
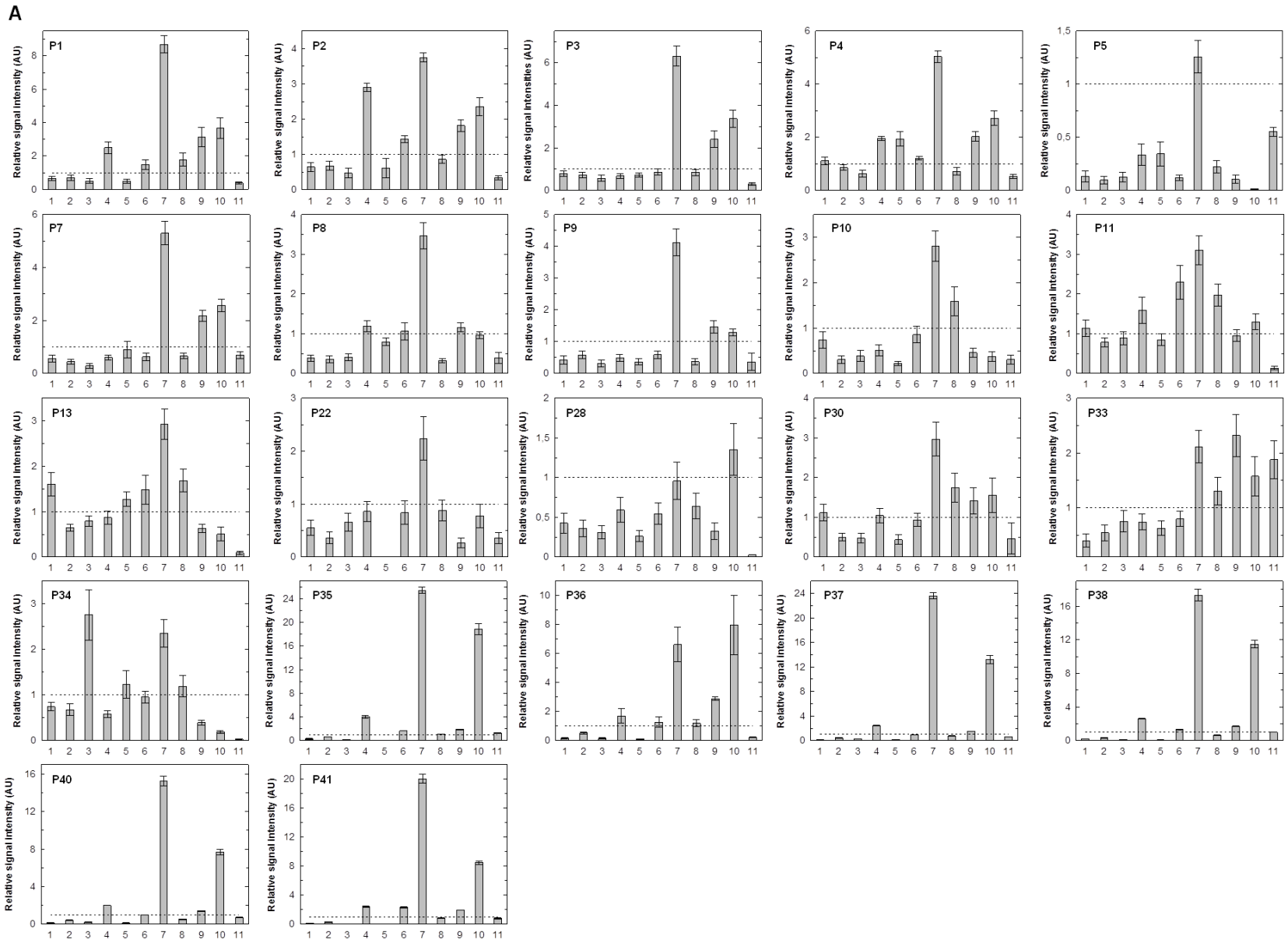


Figure 17: Set-up of the dot blot analysis and serum IgE binding of ΔNCS$_{N42/P49}$ variants.

(A) Dot blot membranes were subdivided into 13 segments of 1x1 cm each. 1 µg of respective ΔNCS$_{N42/P49}$ variant was plotted and binding of serum IgE was detected. (B) Binding of serum IgE to ΔNCS$_{N42/P49}$ variants. Representative examples of the study group and a non-atopic control are shown. Patients' numbers (P11, P20, P26, P30, P33, P36, P37, P41 and P49) are indicated at each blot.

All IgE signals were quantified by densitometry (Figure 18). Each spot was quantified at three different exposure times and mean value as well as standard deviation was calculated. Intensities of each variant were correlated to mean values of $\Delta 51NCS_{N42/P49}$ model protein plus three times standard deviation which was set to 1. IgE signals obtained with study population were identified as positive, if mean values minus standard deviation are above (>1) signal of $\Delta 51NCS_{N42/P49}$ model protein plus three times standard deviation. Results obtained by quantification of IgE binding with all patients are shown in Figure 18.


A
P1
P2
P3
P4
P5
P7
P8
P9
P10
P11
P13
P22
P28
P30
P33
P34
P35
P36
P37
P38
P40
P41
Relative signal intensity (AU)

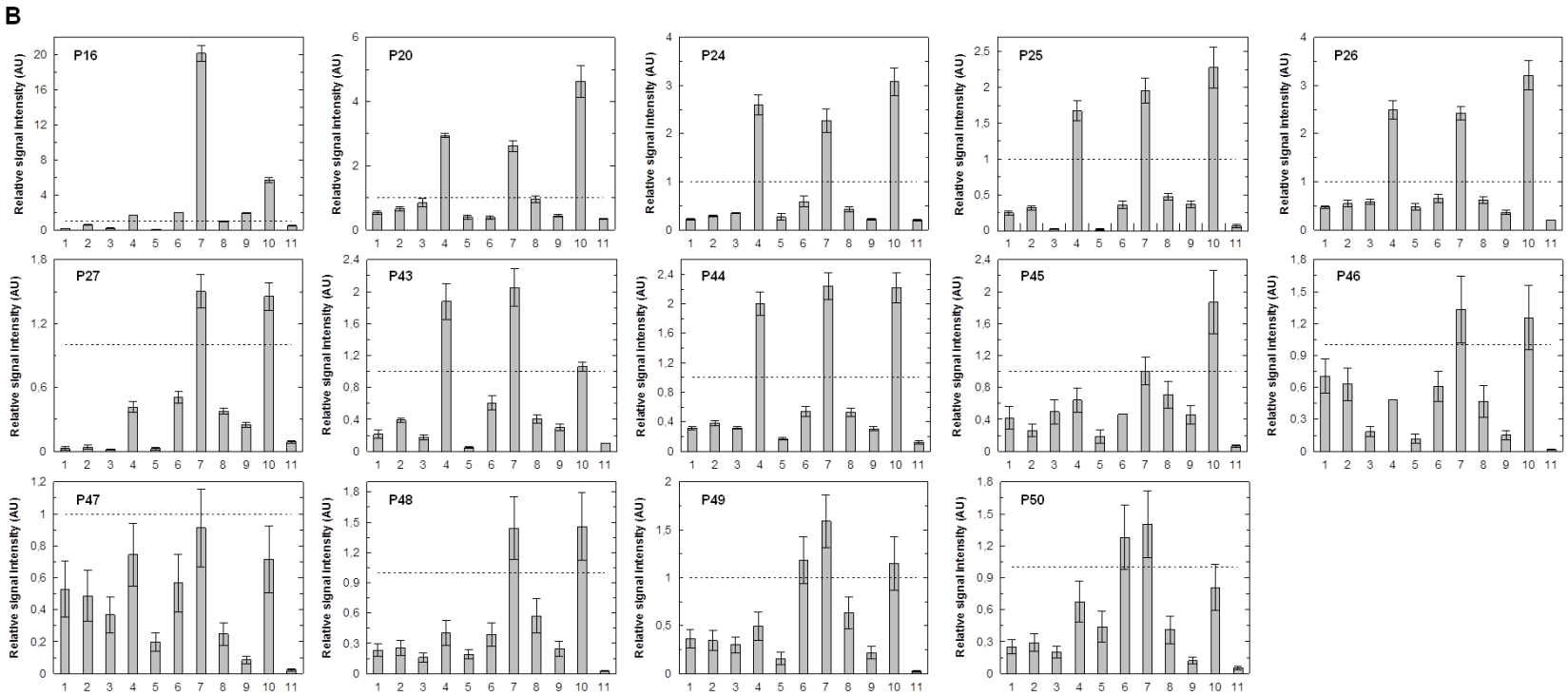


Figure 18: IgE binding to ΔNCS$_{N42/P49}$ variants quantified by densitometry.

Sera of subjects sensitized to Bet v 1 and Gly m 4 with (A, 22 subjects) and without (B, 14 subjects) soybean allergy are shown with relative signal intensities in arbitrary units (AU). Intensities of each dot blot signal were quantified with ImageJ by analysis of three different exposure times. Mean values and standard deviations were correlated to mean values of Δ51NCS$_{N42/P49}$ plus three times standard deviation which was set to 1 (dotted line). Signals were considered as positive, if mean values minus standard deviation are higher than 1. X-axis: 1-11 represent variants ΔNCS$_{N42/P49}$_1-11 introduced above.

Evaluation of IgE signals revealed different IgE binding patterns in both groups. Altogether 39% (14/36) of all sera showed more than three positive IgE signals. Only for one patient of the study population no IgE signal could be identified as positive (patient 47) while serum of subject 35 had IgE against a total number of seven different $\Delta NCS_{N42/P49}$ variants. Mean values for number of identified IgE-binding $\Delta NCS_{N42/P49}$ variants were 3.8 for patients with and 2.2 for patients without soybean allergy. Patients with soybean allergy had identical IgE signals in patients 2, 38 and 41, patients 5 and 22, patients 7 and 9, as well as in patients 36, 37 and 40, with five ($\Delta NCS_{N42/P49}$_4, 6, 7, 9 and 10), one ($\Delta NCS_{N42/P49}$_7), three ($\Delta NCS_{N42/P49}$_7, 9 and 10) and four ($\Delta NCS_{N42/P49}$_4, 7, 9 and 10) IgE-binding $\Delta NCS_{N42/P49}$ variants identified, respectively. Further twelve patients with soybean allergy showed individual IgE-binding patterns. For patients without soybean allergy only four different IgE epitope patterns were identified. Patient 16 showed IgE binding to five $\Delta NCS_{N42/P49}$ variants ($\Delta NCS_{N42/P49}$_4, 6, 7, 9 and 10), while patients 27 and 48, as well as patients 46, 49 and 50 had positive signals with only two ($\Delta NCS_{N42/P49}$_7 and 10) and one $\Delta NCS_{N42/P49}$ variant ($\Delta NCS_{N42/P49}$_7), respectively. For the remaining six patients of this group (patients 20, 24, 25, 26, 43 and 44) same three IgE-binding $\Delta NCS_{N42/P49}$ variants ($\Delta NCS_{N42/P49}$_4, 7 and 9) could be observed. As already shown for immunoblot analysis several weak IgE binding signals can be observed with $\Delta 51NCS_{N42/P49}$ and $\Delta 29NCS_{N42/P49}$ possibly due to identical amino acids between NCS and Gly m 4 or Bet v 1a (see 3.4.2). IgE binding frequencies of all $\Delta NCS_{N42/P49}$ variants regarding both groups of patients are shown in Figure 19.

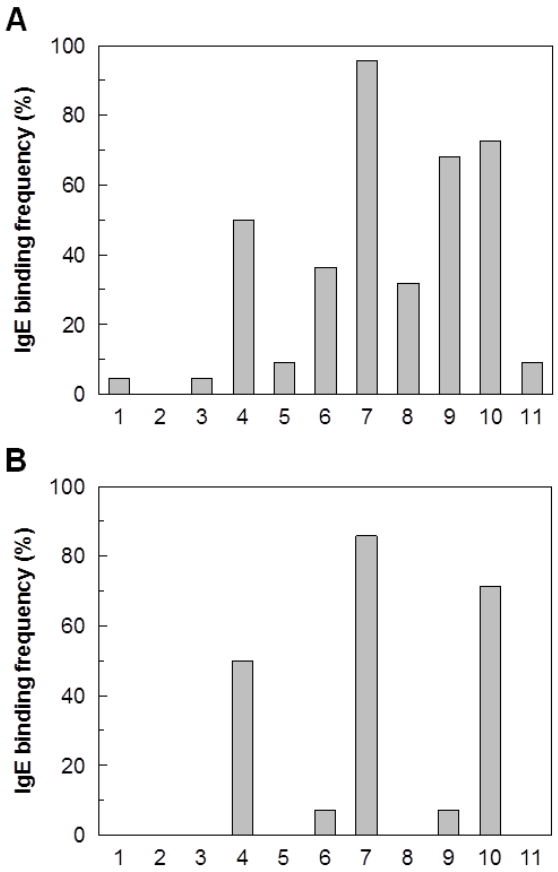


Figure 19: IgE binding frequencies of $\Delta NCS_{N42/P49}$ variants.

IgE binding frequencies of eleven generated $\Delta NCS_{N42/P49}$ variants are shown for patients with (A) and without soybean allergy (B). Frequencies were calculated as percentage of analyzed patients (A: 22, B: 14 subjects, respectively). 1-11 represent variants $\Delta NCS_{N42/P49}$_1-11 introduced above.

IgE binding frequencies differed markedly between subjects with (Figure 19 A) and without (Figure 19 B) soybean allergy. NCS variant identified with highest frequency was $\Delta NCS_{N42/P49}$_7 in both groups of patients with frequencies of 95% (21/22) for patients with and 86% (12/14) for patients without soybean allergy. In addition to $\Delta NCS_{N42/P49}$_7, IgE binding frequencies ≥50% were observed for $\Delta NCS_{N42/P49}$_4 (50%, 11/22), 9 (68%, 15/22) and 10 (73%, 16/22) in patients with and for $\Delta NCS_{N42/P49}$_4 (50%, 7/14) and 10 (71%, 10/14) in patients without soybean allergy. $\Delta NCS_{N42/P49}$_6 and 8 showed IgE binding in 36% (8/22) and 32% (7/22), while residual variants $\Delta NCS_{N42/P49}$_1, 3, 5 and 11 only bound IgE in <10% of patients with soybean allergy. For patients without soybean allergy IgE binding frequencies <10% were identified for $\Delta NCS_{N42/P49}$_6 and 9 while residual six $\Delta NCS_{N42/P49}$ variants showed no IgE signals in any sera of this group. Only $\Delta NCS_{N42/P49}$_2 had no IgE binding with any of used sera. Comparing both groups of patients, major differences were observed with $\Delta NCS_{N42/P49}$_9 showing an IgE binding frequency of 68% in

patients with and only 7% in patients without soybean allergy. $\Delta NCS_{N42/P49}\_9$ revealed a sensitivity of 68% and a specificity of 93% resulting in a positive (PPV) and negative predicted value (NPV) of 94% and 65%, respectively. Therefore residues presented by $\Delta NCS_{N42/P49}\_9$ might be more relevant in the group of subjects with compared to patients without clinically confirmed soybean allergy and $\Delta NCS_{N42/P49}\_9$ might be used as biomarker to differ between soybean allergic patients and subjects only sensitized to Gly m 4. Assembling data of all patients irrespective of their clinical reactivity to soy some differences are observable between majorities of NCS variants. Variants $\Delta NCS_{N42/P49}\_1$, 2, 3, 5, 8 and 11 showed IgE binding with less than 20% of all patients while the remaining variants $\Delta NCS_{N42/P49}\_4$, 6, 7, 9 and 10 bound IgE of 25-92% of all subjects. Summary of identified IgE epitopes with all analyzed patients is shown in Table A 3.

3.5.3 Depletion of IgE binding to Gly m 4

Analysis of eleven $\Delta NCS_{N42/P49}$ variants revealed six putative IgE epitope patches on Gly m 4. To verify if NCS variants and respective epitope patches bind Gly m 4-specific IgE, serum IgE depletion experiments were performed. Briefly, prior to analysis of binding between IgE and rGly m 4, serum was preincubated with rGly m 4, $\Delta 51NCS_{N42/P49}$ or a mixture of $\Delta NCS_{N42/P49}$ variants. Resulting allergen-IgE complexes were precipitated via allergen's His-Tag and serum supernatant was analyzed for Gly m 4-specific IgE binding using ELISA. Results of these depletion experiments are shown in Figure 20.

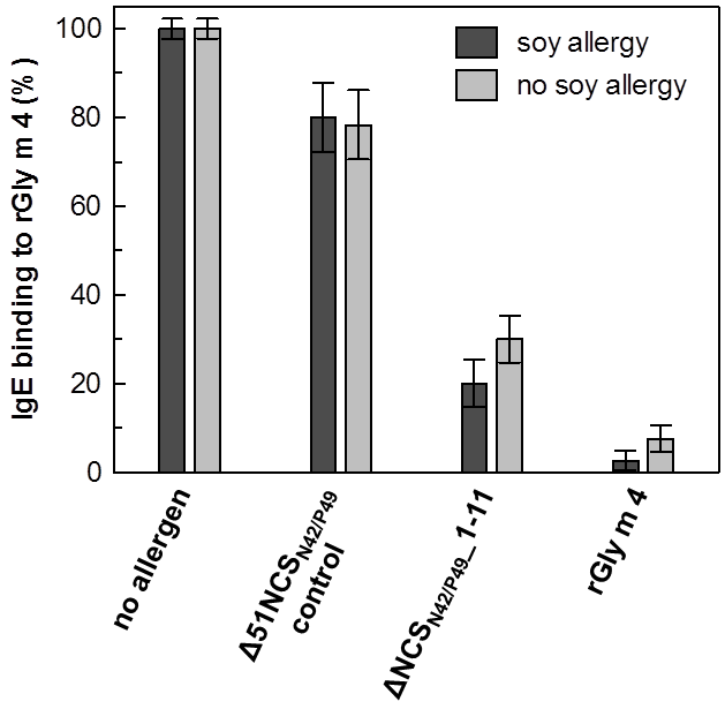


Figure 20: Depletion of IgE binding to rGly m 4.

Pooled sera of patients with (black) and without (grey) birch-related allergy to soy were incubated with rGly m 4 or $\Delta NCS_{N42/P49}$ variants prior to precipitation of the His-tagged proteins via Ni-NTA. Subsequently the supernatants were analyzed for Gly m 4-specific IgE binding. Serum without depleted IgE antibodies shows IgE binding to rGly m 4 and was set to 100%. Percentage of IgE binding to rGly m 4 is shown for incubation with $\Delta 51NCS_{N42/P49}$ control, a mixture of all eleven generated $\Delta NCS_{N42/P49}$ variants and rGly m 4. To present the same number of all IgE epitopes $\Delta 51NCS_{N42/P49}$ and rGly m 4 as well as each variant were used in equimolar amounts (epitope-wise).

Without prior depletion of serum IgE antibodies, a maximum binding between Gly m 4 and IgE was observable with both groups of patients and set to 100%. Irrespective of the clinical reactivity to soy incubation of serum pools with $\Delta 51NCS_{N42/P49}$ caused a slight reduction of Gly m 4-specific IgE binding. In detail a reduction of IgE binding to 80.0 ± 7.7% for patients with soy allergy and 78.3 ± 7.8% for patients without allergy to soybean was observed, respectively. This slight reduction in binding is likely caused by minor IgE interaction of model protein $\Delta 51NCS_{N42/P49}$ with several patients because of identical amino acids in $\Delta 51NCS_{N42/P49}$ and Gly m 4 or Bet v 1a sequence as already shown in 3.4.2 and 3.5.2. In contrast a strongly reduced IgE binding to rGly m 4 was observable when serum IgE was depleted by $\Delta NCS_{N42/P49}$ variants or rGly m 4 itself. By addition of all NCS variants ($\Delta NCS_{N42/P49}$_1-11) representing six distinct IgE epitopes a remaining IgE binding of 20.0 ± 5.3% for the group of patients with soy allergy was determined. In contrast subjects sensitized but not clinically soy allergic showed a reduction in IgE binding to 30 ± 5.3% with eleven $\Delta NCS_{N42/P49}$ variants. Both reductions indicate the Gly m 4 specificity of the identified interactions

between $\Delta NCS_{N42/P49}$ variants and serum IgE antibodies in both groups of patients. Depletion of IgE with rGly m 4 caused an even increased reduction in IgE binding as shown for $\Delta NCS_{N42/P49}$ variants. In detail, IgE binding to rGly m 4 was reduced to almost 2.7 ± 2.1% caused by prior depletion with the same allergen using pool of sera from patients showing soybean allergy. A comparable reduction to 7.6 ± 3.0% in IgE binding to rGly m 4 was observable by Gly m 4 depletion with patients sensitized but without soy allergy. The observed differences in IgE binding between rGly m 4 and all $\Delta NCS_{N42/P49}$ variants might indicate the presence of further IgE epitope(s) not identified yet or might reflect a decreased affinity in binding between IgE and $\Delta NCS_{N42/P49}$ variants compared to IgE and rGly m 4 in both groups.

3.6 Analysis of putative functional IgE epitopes of rGly m 4

3.6.1 Generation and characterization of rGly m 4

Recombinant Gly m 4 was expressed with cleavable N-terminal His-tag in *E. coli* as both, soluble and insoluble (inclusion bodies) protein. rGly m 4 was purified to homogeneity (see 2.4.4) as soluble protein with a total protein yield of about 20 mg/L bacterial suspension (Figure 21).

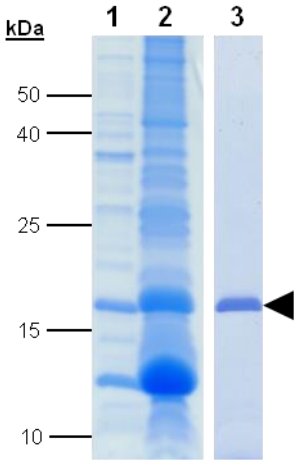


Figure 21: Purity of recombinant Gly m 4.

Purified rGly m 4 after small-scale expression in *E. coli* analyzed in Coomassie-stained gels upon SDS-PAGE. Shown are soluble fraction (lane 1) and inclusion bodies (2) obtained by cell lysis using BugBuster extraction kit and purified rGly m 4 (3). For BugBuster analysis, loaded volume of inclusion body preparation was 3 times the volume of soluble fraction. Arrow head (◄) marks rGly m 4 protein.

To analyze structural integrity of recombinant Gly m 4, secondary and tertiary structure was determined by CD and 1D-^{1}H-NMR spectroscopy (Figure 22).

Furthermore potential protein oligomerisation/aggregation was addressed by dynamic light scattering (DLS).

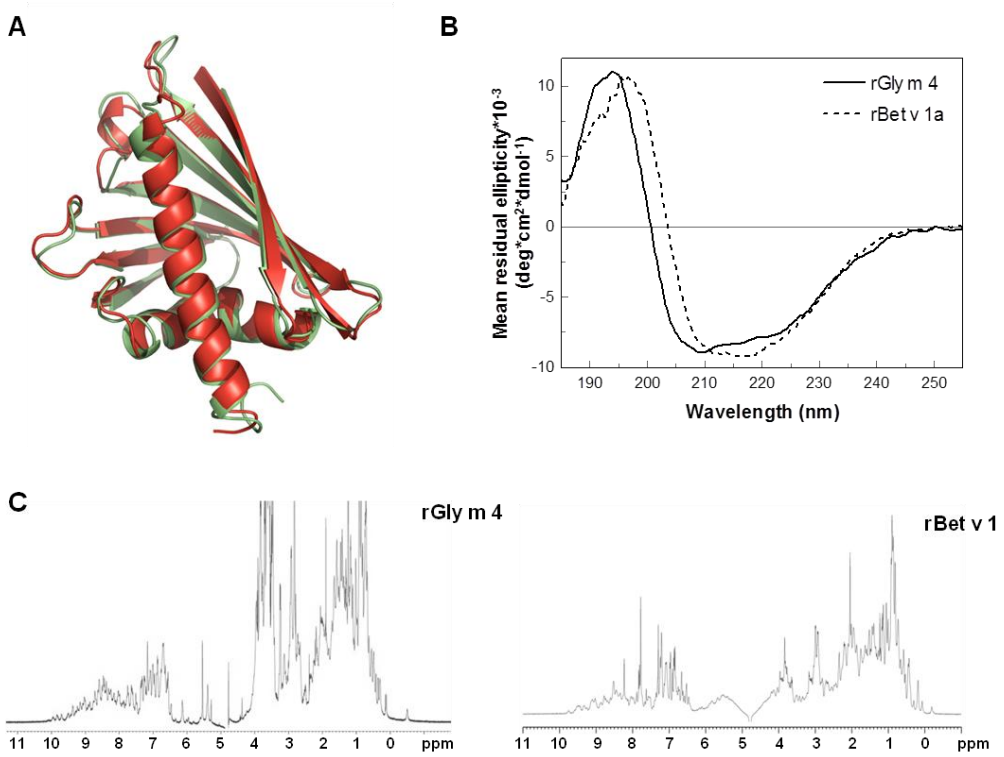


Figure 22: Secondary and tertiary structure of rGly m 4 and rBet v 1a.

(A) Overlay of secondary structure topologies of rGly m 4 (red) and rBet v 1a (green) generated with PyMOL. (B) Circular dichroism of rGly m 4 and rBet v 1a (10 µM each). (C) 1D-^{1}H-NMR of 40 µM of rGly m 4 (left panel) and 30 µM rBet v 1a (right panel).

rGly m 4 had typical Bet v 1-type conformation with a high content of both β-sheets and α-helices as determined in circular dichroism. With 1D-^{1}H-NMR spectroscopy the native Gly m 4 protein conformation was confirmed, as the NMR signals indicated a high degree of resonance dispersion at 8.5-6.5 ppm and 4.5-0.5 ppm characteristic for amid protons and α-protons, as well as protons of methyl groups resulting from secondary structure elements. Additional signals are present in both rBet v 1a and rGly m 4 below 0.5 ppm, above 8.5 ppm and between 5.0-6.5 ppm and only appear in well-folded proteins with existing secondary and tertiary structure. Dynamic light scattering analysis revealed a hydrodynamic radius (R_H) of 2.81 ± 0.27 nm (rGly m 4) which was slightly higher compared to R_H (rBet v 1a) of 2.49 ± 0.39 nm (Ferreira *et al.*, 1993). Differences might result from additional amino acids in Gly m 4 sequence representing linker and adjacent His_6-tag residues causing an increased hydrodynamic radius. Nevertheless R_H indicated a monodisperse distribution of rGly m 4 as monomeric protein in solution with no aggregates detectable.

3.6.2 Characterization of substitutional rGly m 4 variants

To test the impact of amino acids on Gly m 4-specific IgE binding identified with the NCS model system (3.2.2) several multiple substitutional rGly m 4 variants were generated (Table 16).

Table 16: Generated variants of rGly m 4.

Name	Number of substitutions	Amino acid substitutions
rGly m 4_4x	4	T93, E95, K96, E120
rGly m 4_9x	9	E6, K32, S39, D60, T93, E95, K96, E120, E126
rGly m 4_10x	10	E6, D27, K32, D35, S39, D60, D92, E95, K118, E126
rGly m 4_11x	11	E6, D27, K32, D35, S39, D60, D92, E95, K118, E126, Y149
rGly m 4_14x	14	E6, D27, K32, D35, S39, E44S, D60, N77K, D92, E95, N108K, K118, E126, K133E

Substituted amino acids of each variant are shown. If not stated otherwise, amino acids were changed to alanine. Amino acids are colored according to their predicted epitope assignment (see Figure 13).

All of the substituted residues correlate to the putative IgE epitopes identified with NCS model protein analysis (3.4.2) and are shown in Figure 23.

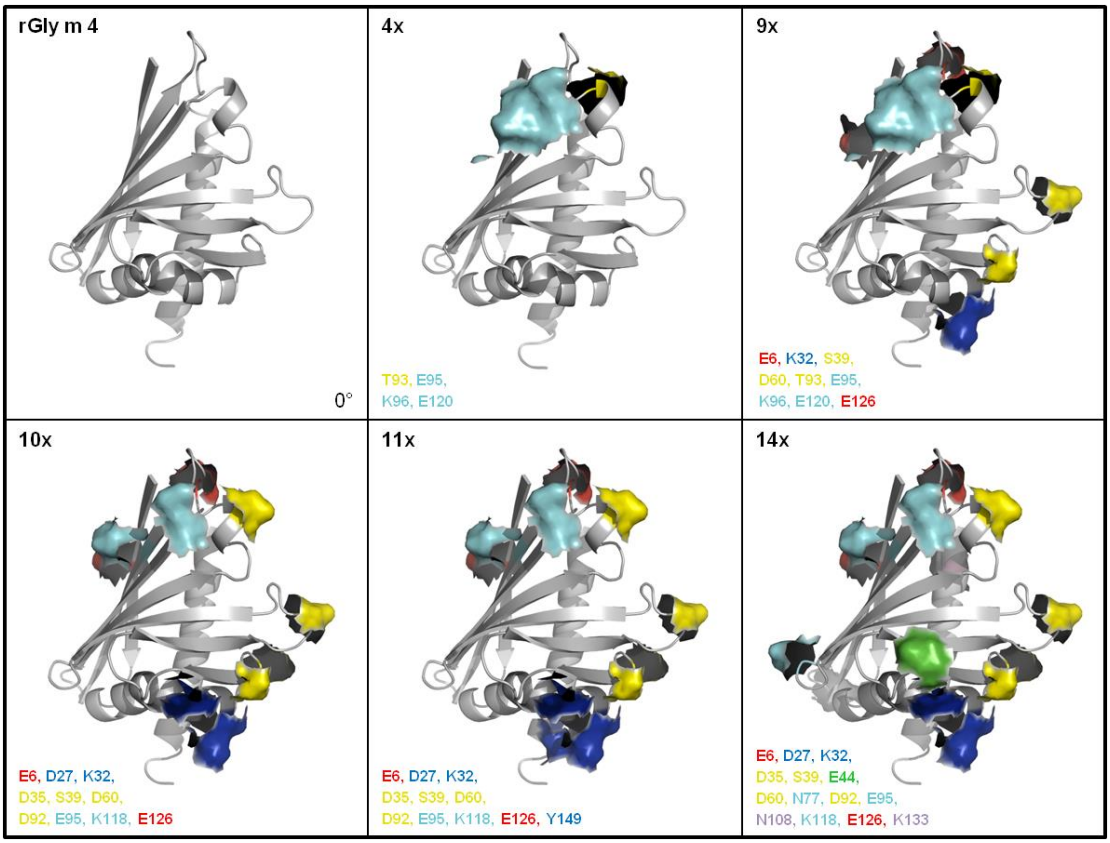


Figure 23: Secondary structure topologies of rGly m 4 and substitutional variants.

Shown are secondary structural elements and substituted residues of 4x (blue), 9x (green), 10x (red), 11x (yellow) and 14x (purple) Gly m 4 variants as well as their corresponding molecular surface. Molecules were generated with PyMOL and pdb code 2K7H with front view (0°) shown.

Amino acids E6 and E126 correspond to epitope I, N108 and K133 to epitope II, D27 and K32 are parts of epitope III, D35, S39, D60, D92 and T93 correspond to epitope IV, E44 to epitope V and K118 and E120 to epitope VI, respectively. Amino acids E44, N77, N108 and K133 were already identified as parts of IgE-binding epitopes in Holm *et al.* and Spangfort *et al.*, where respective substitutions in Bet v 1 resulted in a reduced binding of serum IgE (Spangfort *et al.*, 2003; Holm *et al.*, 2004). Residues E95 and K96 were identified as potential IgE-binding in previous Δ51NCS analyses (Nürnberg, 2014, Master thesis) and Y149 belongs to different mimotopes identified by Mittag *et al.* (Mittag *et al.*, 2006). N77, E95, K96 could be assigned to epitope VI and Y149 to epitope III, respectively. To analyze the influence of substitutions on IgE binding the rGly m 4 variants were purified (Figure 24 A) and secondary structure content analyzed by CD spectroscopy (Figure 24 B).

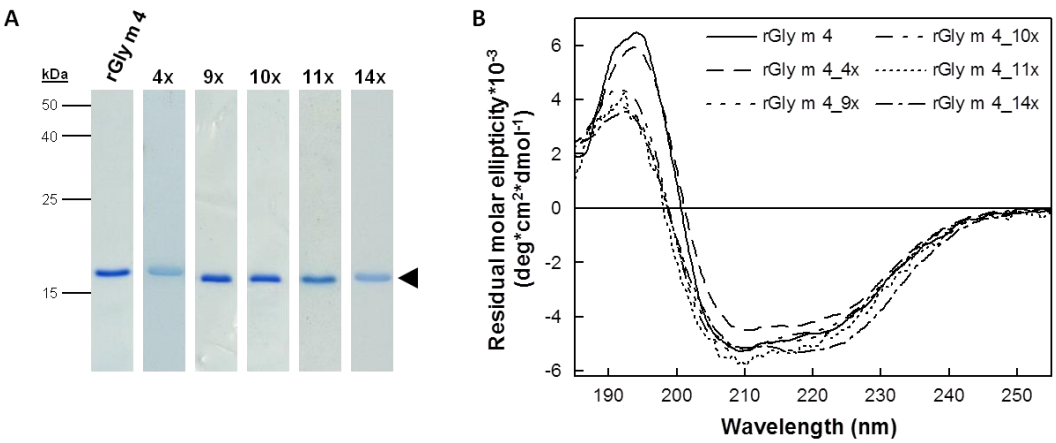


Figure 24: Purity and circular dichroism of substitutional rGly m 4 variants.

(A) Purified rGly m 4 and substitutional variants were analyzed in Coomassie-stained gels upon SDS-PAGE. Arrow head (◀) mark recombinant target proteins. (B) CD spectra of rGly m 4 and 4x, 9x, 10x, 11x and 14x variant recorded at 5 µM each. Spectra are presented as mean residual ellipticity*10^{-3} at a given wavelength and baseline corrected.

Furthermore amino acid sequences of rGly m 4 and rGly m 4_11x variant were analyzed according to LC-MS analyses (see 2.4.19) and MS data confirmed identities of both proteins (Table 17). Details of peptides detected by MS verifying relevant amino acid substitutions of rGly m 4_11x variant are shown in Figure A 1 in the appendix.

Table 17: Summary of the MS data of the confirmation of rGly m 4 and rGly m 4_11x variant.

Protein	Acc.No.*	Score[#]	Peptides[§]	Coverage[$]	Mass Error[¨]
rGly m 4	PEI015	16135	16	74%	4.2ppm
rGly m 4_11x	PEI014	14715	17	95%	3.9ppm

Listed are results of LC-MS analyses for confirmation of amino acid sequence of rGly m 4 and rGly m 4_11x with * internal accession number; [#] protein score by PLGS; [§] number of tryptic peptides detected by MS; [$] Percentage of sequence covered by MS; " average precursor mass error.

Since all rGly m 4 variants exhibited Gly m 4-type circular dichroism spectra, the proteins were considered suitable for IgE binding analyses as described in the following chapter.

3.6.3 Analysis of IgE binding of substitutional variants of rGly m 4

Amino acid substitutions of rGly m 4 were analyzed for their relevance in IgE binding with immunoblot and ELISA inhibition. Reduced IgE binding to rGly m 4 variants as compared to wild type rGly m 4 would suggest contribution of substituted amino acids in interaction with the IgE antibody. To identify such functional IgE epitopes several sera of patients with (group 1) and without (group 2) birch-related allergy to soy were tested. First, rGly m 4_11x variant was analyzed for reduced IgE binding in comparison to wild type rGly m 4 with a total number of 16 patients of which twelve were allergic to soy and 4 were tolerant. Therefore immunoblot analysis was performed with both proteins incubated in parallel with patients' sera (Figure 25).

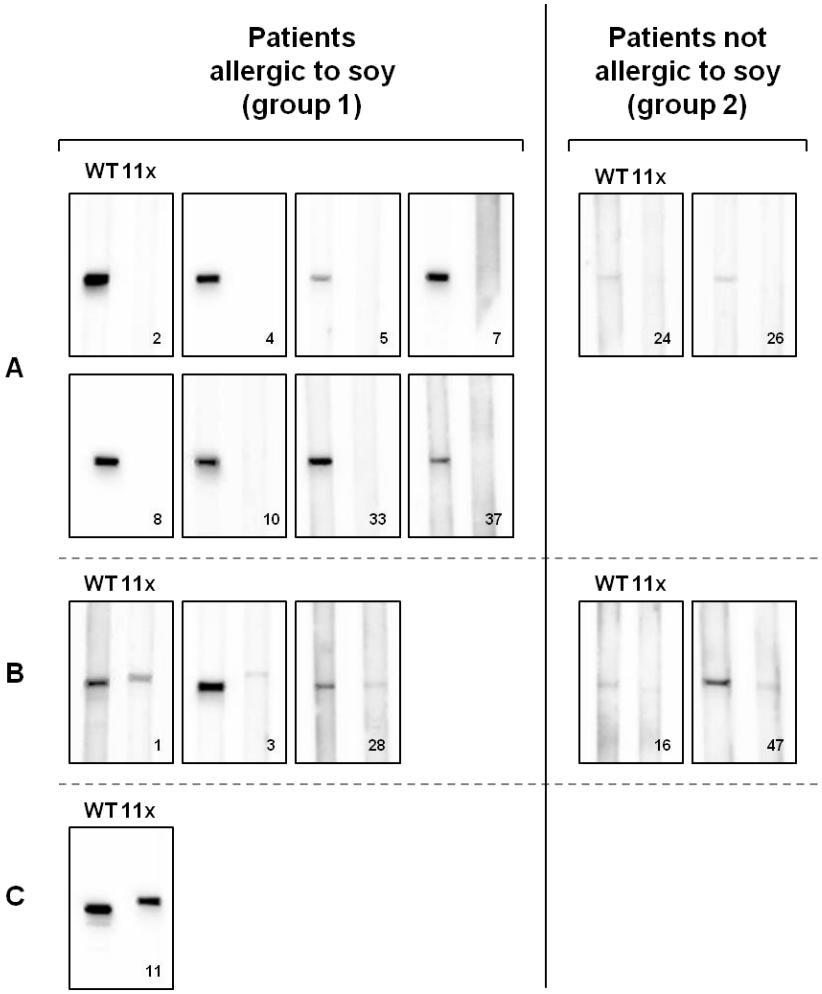


Figure 25: IgE binding of rGly m 4 versus rGly m 4_11x.

Serum IgE binding of 12 patients allergic (group 1, left panel) and 4 patients not allergic to soy (group 2, right panel) was analyzed with rGly m 4 and rGly m 4_11x. Immunoblot results were divided into subjects with no (A), minor (B) and strong (C) binding of serum IgE to rGly m 4_11x compared to wild type allergen. At each blot the specific serum number is indicated.

Analyzed patients were classified according to their IgE binding with rGly m 4_11x. Strong reduction in IgE binding to the rGly m 4 variant was observed in 63% (10/16) of all patients. Regarding both groups analyzed, 67% (8/12) of patients with and 50% (2/4) of patients without soybean allergy, respectively, showed a strong reduction in IgE binding with rGly m 4_11x. In comparison, a clearly reduced IgE binding to the rGly m 4 variant was observed in 31% (5/16) of all patients, 25% (3/12) of group 1 and 50% (2/4) of group 2, respectively. A comparable IgE binding of both studied rGly m 4 proteins was only observable in one patient (8%) with clinical allergy to soy. These results indicate the impact of substituted amino acids on IgE binding. At least one IgE epitope was affected in several patients through amino acid substitutions causing reduced binding of IgE antibodies compared to wild type rGly m 4. Nevertheless, remaining IgE binding to rGly m 4_11x especially in patient 11 suggested the presence of further epitopes of Gly m 4. To get a more detailed view on influence of substituted residues for IgE binding several sera were tested with rGly m 4_4x, 9x, 10x and 14x. Because of their overlapping substitutions it might be possible to localize important residues for IgE binding. Additional analyses were performed in solution using ELISA inhibition. Again, same 16 patients of group 1 and 2 were analyzed with results shown in Figure 26.

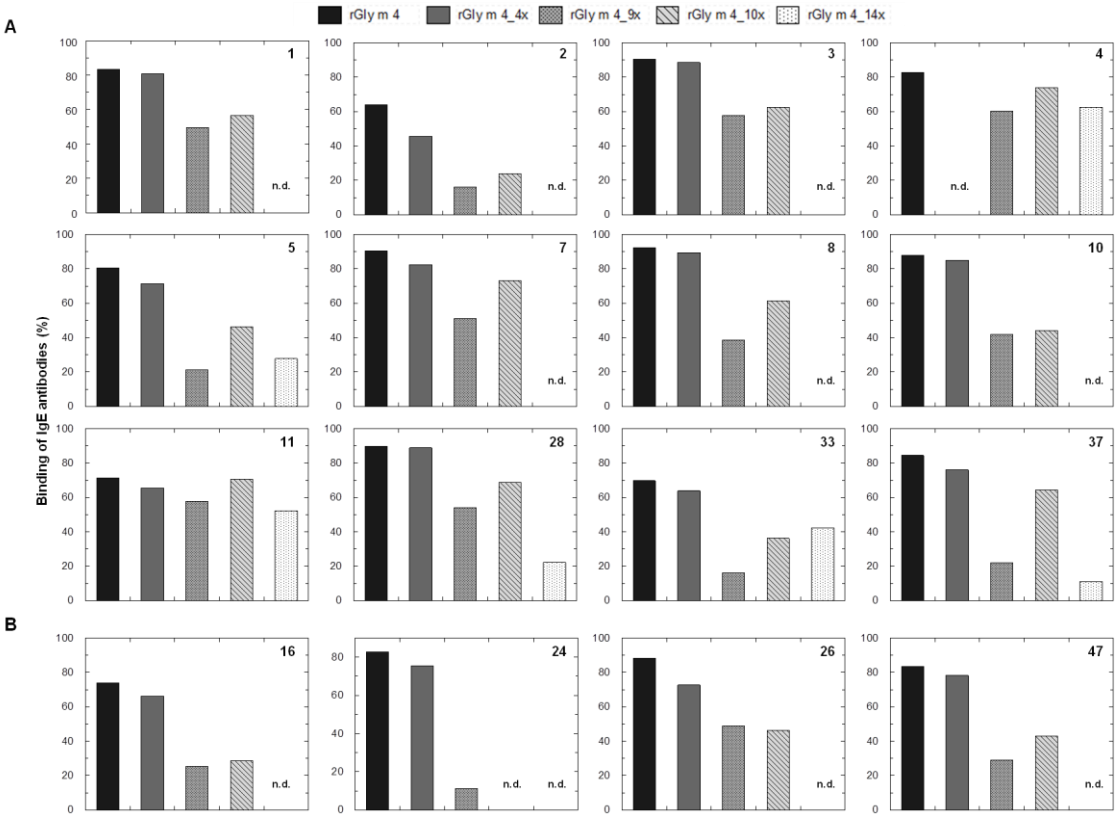


Figure 26: Reduced IgE binding to rGly m 4 in presence of recombinant substitutional Gly m 4 variants.

Binding of serum IgE to 200 ng rGly m 4 was inhibited with 500 ng rGly m 4, rGly m 4_4x, 9x, 10x or 14x variant in ELISA. Sera of 12 subjects allergic (A) and 4 subjects not allergic to soybean (B) were used. Number of each patient is indicated. The average of two measurements is shown. n.d.: not determined.

ELISA analysis revealed different IgE binding patterns with rGly m 4 variants in both groups of patients irrespective of their allergic symptoms. As control IgE binding to rGly m 4 can be inhibited with rGly m 4 to 63.9-92.3%. For all rGly m 4 proteins only a single-point measurement using 2.5 time excess of inhibitor compared to rGly m 4 present on ELISA plate was performed. Comparable results to rGly m 4 were obtained with rGly m 4_4x variant showing inhibition to 45.6-92.1% of IgE binding in analyzed subjects. In contrast stronger inhibition of IgE binding was observed with rGly m 4_9x, 10x and 14x variant. rGly m 4_9x which is based on substitutions present in rGly m 4_4x variant inhibited IgE binding to 16.2-60.3% indicating the importance of additional substitutions for binding of IgE antibodies in several sera (No. 5, 8, 10, 33, 37, 16, 24, 47). With a inhibition of IgE-allergen binding to 23.8-73.9% rGly m 4_10x is comparable to rGly m 4_9x which both show six shared substitutions. rGly m 4_14x which was tested only in six patients of group 1 exhibited four additional substitutions compared to rGly m 4_10x and inhibited IgE binding to 11.0-62.6%. Interestingly, in some patients (No. 5, 28, 37) rGly m 4_14x showed a

lower inhibition level compared to rGly m 4_10x. In patient 11 only small differences with all rGly m 4 variants was observable showing IgE binding inhibited from 52.3% with rGly m 4_14x to 71.2% with rGly m 4. Overall, a patient-specific pattern of inhibition was observed in both groups of patients analyzed. Summary of results determined by inhibition ELISA experiments are listed in Table 18.

Table 18: Inhibition of IgE binding to rGly m 4 with substitutional rGly m 4 variants.

A	Decrease in IgE binding (%) with substitutional variants compared to wild type rGly m 4 as inhibitor			
Variant	**No. of analyzed patients**	**≥40%**	**20-40%**	**<20%**
4x	12	0	0	12 (100%)
9x	12	6 (50%)	5 (42%)	1 (8%)
10x	12	2 (17%)	7 (58%)	3 (25%)
14x	6	3 (50%)	2 (33%)	1 (17%)

B	Decrease in IgE binding (%) with substitutional variants compared to wild type rGly m 4 as inhibitor			
Variant	**No. of analyzed patients**	**≥40%**	**20-40%**	**<20%**
4x	4	0	0	4 (100%)
9x	4	3 (75%)	1 (25%)	0
10x	4	3 (75%)	0	1 (25%)
14x	0	-	-	-

Decrease in IgE binding in % with rGly m 4_4x, 9x, 10x and 14x variants compared to inhibition determined with wild type rGly m 4 as inhibitor. Values are categorized in strong (≥40%), medium (20-40%) and weak (<20%) reduction of IgE binding. Sera of 12 subjects sensitized and soy allergic (A) and 4 subjects sensitized but without clinical soybean allergy (B) were used.

Comparing the decrease in IgE binding obtained with different rGly m 4 variants as inhibitors in correlation with rGly m 4 wild type protein an increased number of substitutions comes along with a considerable decrease in IgE binding in most of the patients. Especially rGly m 4_9x, 10x and 14x variants showed reduced binding to

serum IgE directed against rGly m 4 in all patients analyzed. Therefore amino acids substituted in these variants seemed to be critical for binding IgE antibodies and thus might be part of several IgE epitopes. Nevertheless there are still sera of patients observable in both groups showing only slightly decreased IgE binding even with rGly m 4_10x or 14x as inhibitors indicating the presence of further IgE epitopes left on Gly m 4 not influenced by introduced substitutions. A detailed analysis of which residue actually contributes to antibody binding is not possible but first hints on putative functional epitopes can be concluded. For example with patient 5 a strong decrease in IgE binding with rGly m 4_9x (21.4%) compared to rGly m 4_4x (71.3%) which differs in residues E6, K32, S39, D60 and E126 was observed. Therefore at least one of these residues might be important for serum IgE binding in this subject. Same holds true for patients 10, 33 and 37 of group 1 as well as patients 16, 34 and 47 of group 2. Furthermore a considerably decreased IgE binding was observed with rGly m 4_14x (11%) compared to rGly m 4_10x (64.2%) in patient 37. Both variants only differ in additional substitutions E44S, N77K, N108K and K133E introduced in rGly m 4_14x indicating their potential relevance for binding IgE antibodies. As shown for patient 37 these differences were also observed in patient 28. Despite the missing results of rGly m 4_14x in patients without soybean allergy no characteristic differences of one group of patients could be determined. Therefore no differences in the epitope profiles between both groups of patients were observed with substitutional rGly m 4 variants. In sum, the analysis of the substitutional variants of Gly m 4 revealed 18 potential functional epitopes, i. e. amino acids which might dominate the energetics of allergen-IgE binding, as defined in chapter 3.3.

3.7 Analyzing IgE binding avidity with rGly m 4

For a thorough analysis of potential differences between patients with and without clinical allergy to soy, IgE antibody avidity needs to be addressed. A higher avidity of Gly m 4-specific IgE in patients with allergy compared to patients sensitized but tolerant to soybean could be a potential explanation for the presence of allergic reactions. Therefore IgE binding avidity between rGly m 4 and two pools of patients representing sera of group 1 and 2, respectively, was determined by inhibition ELISA. According to previous results where no differences in serum IgE level was observed in both groups of patients (see 3.1), inhibition of IgE binding to rGly m 4 in both pools

of sera should show comparable results indicating no differences in IgE binding avidity with rGly m 4. Results of inhibition ELISA are shown in Figure 27.

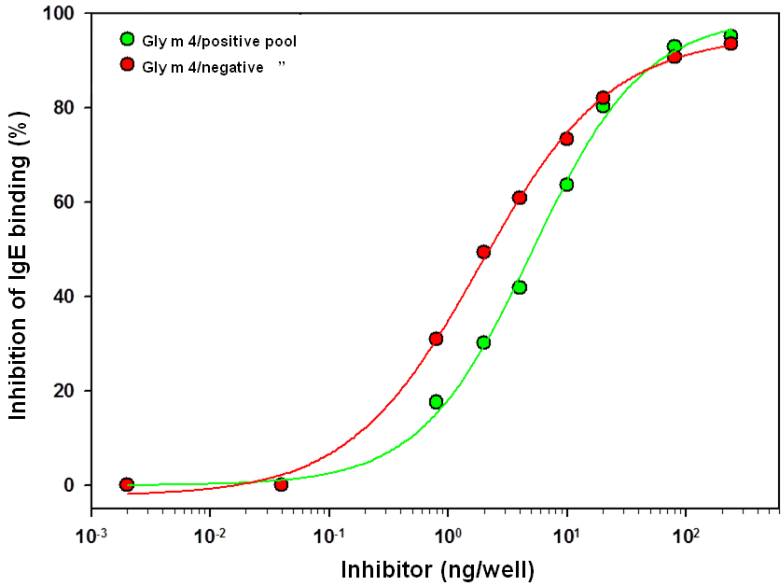


Figure 27: Inhibition of IgE binding to immobilized rGly m 4 with rGly m 4 as inhibitor.

Binding of IgE antibodies of serum pools generated from 27 subjects with (group 1, positive pool – green) and 20 subjects without (group 2, negative pool – red) soybean allergy. IgE binding to immobilized rGly m 4 was inhibited by increasing concentrations of rGly m 4 in ELISA. All curves were fitted by a 4-parameter, logistic sigmoidal curve fit.

With both serum pools a comparable inhibition curve with rGly m 4 as inhibitor was observable. EC_{50} values, the Gly m 4 concentrations needed for half-maximum inhibition of IgE-Gly m 4 binding for positive and negative pool were 5.06 ± 0.47 ng and 1.84 ± 0.14 ng for rGly m 4. With rGly m 4 the pool of negative sera revealed a slightly decreased EC_{50} value compared to the pool generated from positive patients. Overall no differences in avidities regarding both analyzed serum pools with rGly m 4 could be detected which did not support the hypothesis of a potential importance of avidity comparing patients with and without allergy to soy at least in this study population.

3.8 Misfolded allergen variants

3.8.1 Generation of a misfolded variant of Gly m 4 and Bet v 1a

It is widely accepted that IgE-mediated allergy to birch pollen is caused by IgE binding to conformational epitopes of Bet v 1, the major allergen of birch. In birch-associated food allergy however, allergic symptoms may arise up to two hours after ingestion of the allergenic food raising the question whether allergen processing through the gastrointestinal tract and loss of native folding in general might modulate allergen protein conformation and thus IgE binding capability (Bohle *et al.*, 2006). To analyze whether native protein conformation is required for serum IgE-binding in the study population analyzed here, two variants of each rBet v 1a and rGly m 4 lacking the typical Bet v 1-type fold were generated and characterized for physicochemical and immunobiological properties. In detail, two proline residues were introduced into both allergens by site-directed mutagenesis. First substitution was serine at position 112 in Bet v 1a and corresponding residue S111 in Gly m 4 and was already shown to cause a significant change in protein folding for Bet v 1a (Son *et al.*, 1999). In addition, one residue of long C-terminal α-helix was changed to disrupt this secondary element. For Bet v 1a arginine 145 and for Gly m 4 leucine 150 were each substituted to a proline residue. Resulting variants rBet v $1a_{S112P/R145P}$ and rGly m $4_{S111P/L150P}$ are shown in Figure 28.

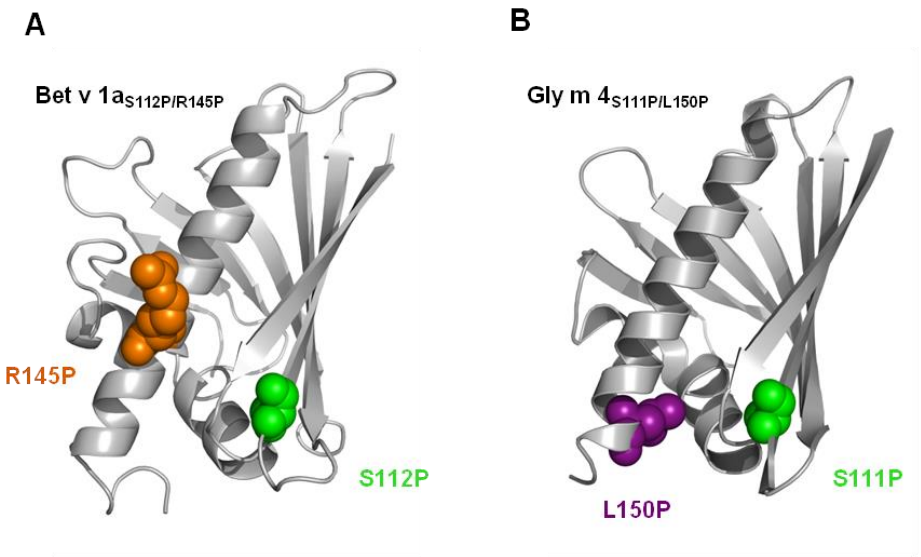


Figure 28: Secondary structure topology of Bet v 1a (A, pdb: 1BV1) and Gly m 4 (B, pdb: 2K7H).

Amino acids S112 (green) and R145 (orange) of Bet v 1a as well as S111 (green) and L150 (purple) of Gly m 4 were exchanged for proline, respectively.

Both variants were heterologously expressed in *E. coli* BL21(DE3) and purified from bacterial inclusion bodies with N-terminal His-tag and affinity chromatography. For comparability reasons rBet v 1a, which was expressed in both, inclusion bodies and soluble fraction, was also purified from *E. coli* inclusion bodies. Only rGly m 4 was derived from soluble fraction as mentioned above (see 3.6.1). In Figure 29 SDS-PAGE with both fractions after cell expression, as well as purified allergens are shown.

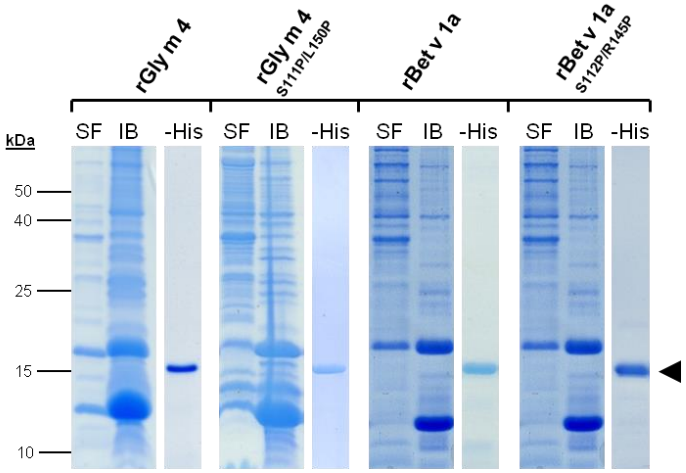


Figure 29: Expression of misfolded variants rGly m 4$_{S111P/L150P}$ and rBet v 1a$_{S112P/R145P}$ in *E. coli*.

Coomassie-stained gel of bacterial lysates of *E. coli* expressing rGly m 4 and rBet v 1a as well as their respective misfolded variants rGly m 4$_{S111P/L150P}$ and rBet v 1a$_{S112P/R145P}$ and purified target proteins analyzed by SDS-PAGE. Soluble fractions (SF) and inclusion bodies (IB) obtained by cell lysis using BugBuster extraction kit are shown. Arrow heads (◄) mark purified recombinant target proteins with N-terminal His-tag removed by thrombin cleavage (-His).

Purified proteins and their respective variants were furthermore treated with thrombin to cleave off the N-terminal His-tag (-His). This caused a reduction in molecular weight from 18.8 kDa to 17.1 kDa for proteins expressed in pET15b vector. Overall the recombinant allergens were highly pure with only minor degradation products of rBet v 1a and rBet v 1a$_{S112P/R145P}$. Using LC-MS/MS the amino acid sequences of both misfolded variants could be confirmed with a sequence coverage of 78% (score 7985) for rBet v 1a$_{S112P/R145P}$ and 86% (score 17230) for rGly m 4$_{S111P/L150P}$, respectively. Results are shown in Table 19. Details of peptides detected by MS verifying relevant amino acid substitutions of rBet v 1a$_{S112P/R145P}$ and rGly m 4$_{S111P/L150P}$ variant are shown in Figure A 1 in the appendix.

Table 19: Summary of the MS data of the confirmation of rBet v 1a$_{S112P/R145P}$ and rGly m 4$_{S111P/L150P}$.

Protein	Acc.No.*	Score#	Peptides§	Coverage$	Mass Error"
rBet v 1a	PEI037	8297	14	72%	8.3 ppm
rBet v 1a$_{S112P/R145P}$	PEI043	7985	13	78%	3.4 ppm
rGly m 4	PEI015	25161	14	72%	4.4 ppm
rGly m 4$_{S111P/L150P}$	PEI059	17230	18	86%	5.7 ppm

Listed are results of LC-MS analyses for confirmation of amino acid sequence of rBet v 1a, rBet v 1a$_{S112P/R145P}$, rGly m 4 and rGly m 4$_{S111P/L150P}$ with * internal accession number; # protein score by PLGS; § number of tryptic peptides detected by MS; $ Percentage of sequence covered by MS; " average precursor mass error.

3.8.2 Characterization of rBet v 1a$_{S112P/R145P}$ and rGly m 4$_{S111P/L150P}$

For a detailed analysis of both, rBet v 1a$_{S112P/R145P}$ and rGly m 4$_{S111P/L150P}$, physicochemical as well as immunological characteristics were evaluated. Firstly, secondary and tertiary structures of each allergen were analyzed by CD and 1D-^{1}H-NMR spectroscopy shown in Figure 30.

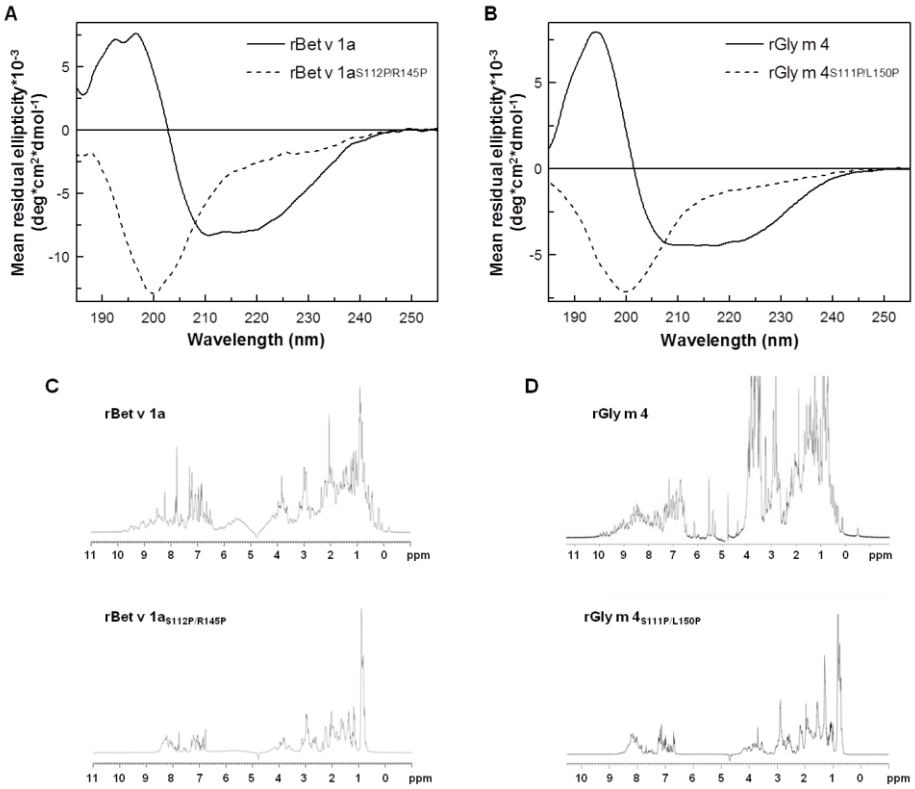


Figure 30: Circular dichroism and 1D-^{1}H-NMR spectra of rBet v 1a$_{S112P/R145P}$ and rGly m 4$_{S111P/L150P}$.

Circular dichroism of rBet v 1a and rBet v 1a$_{S111P/R145P}$ (A) as well as rGly m 4 and rGly m 4$_{S112P/L150P}$ (B). 1D-^{1}H-NMR spectra of 30 µM of rBet v 1a and rBet v 1a$_{S112P/R145P}$ (C) and 40 µM of rGly m 4 and rGly m 4$_{S112P/L150P}$ (D).

Both allergens rBet v 1a and rGly m 4 showed typical CD spectra with a high content of α-helices and β-sheets as already shown in Figure 22. In contrast respective variants indicated different secondary structures. Both rBet v $1a_{S112P/R145P}$ and rGly m $4_{S111P/L150P}$ showed negative mean residual ellipticities between 240 and 185 nm with a maximal negative ellipticity at 200 nm, characteristic for unstructured proteins. Consistent results were obtained when the tertiary structures of the recombinant allergens were analyzed by NMR spectroscopy. rBet v 1a and rGly m 4 had spectra with a high degree of resonance dispersion resulting from amid (6-10 ppm) as well as α- (3.5-6 ppm) and methyl protons (-0.5-1.5 ppm) indicating a characteristic Bet v 1-type protein fold for both allergens. Furthermore these proteins showed well-separated peaks in overall spectra. In contrast, the spectra recorded with rBet v $1a_{S112P/R145P}$ and rGly m $4_{S111P/L150P}$ exhibited less peaks with reduced dispersion especially in the amide region and no signals from 5.0-6.5 ppm and below 0.5 ppm which is characteristic for misfolded proteins or unstructured polypeptides (Page *et al.*, 2005). Therefore substitutions of S112 and R145 (Bet v 1a) or S111 and L150 (Gly m 4) to proline residues disrupted the native protein folds of Bet v 1a and Gly m 4, respectively. For analysis of potential protein aggregation, dynamic light scattering (DLS) was used to determine the hydrodynamic radii of rBet v 1a and rGly m 4 as well as their respective misfolded variants. As already stated in 3.6.1 rBet v 1a and rGly m 4 had comparable R_H of 2.81 ± 0.27 nm (rGly m 4) and 2.49 ± 0.39 nm (rBet v 1a), respectively. In contrast increased hydrodynamic radii of 3.10 ± 0.56 nm for rBet v $1a_{S112P/R145P}$ and 3.54 ± 0.46 nm for rGly m $4_{S111P/L150P}$ were observed, suggesting a more unfolded protein conformation as compared to the wild type proteins. The hydrodynamic radii indicated that the proteins did not aggregate and were mono-disperse in solution.

3.8.3 IgE binding of misfolded variants

According to inhibition ELISA analyses with rGly m 4 shown in Figure 27, IgE binding avidity was furthermore determined between rGly m $4_{S111P/L150P}$ and the two pools of patients with and without clinical allergy to soy, respectively. Results of inhibition ELISA with rGly m $4_{S111P/L150P}$ is shown in Figure 31.

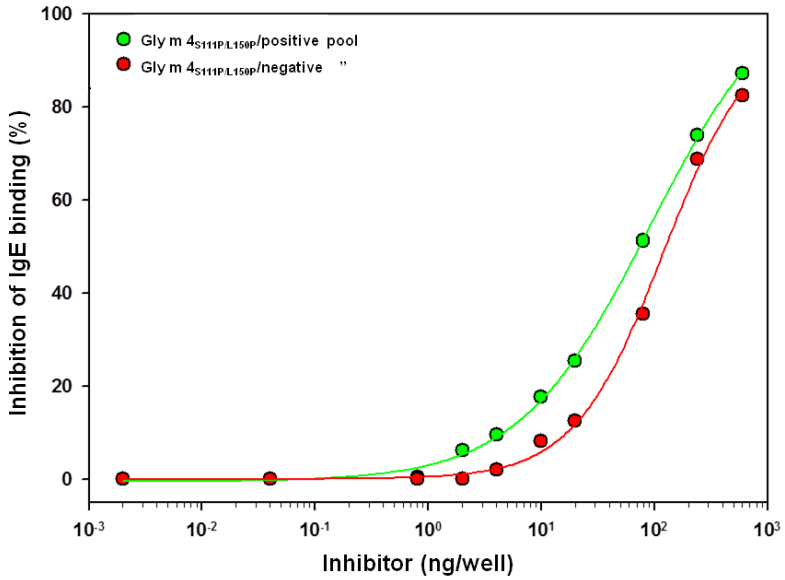


Figure 31: Inhibition of IgE binding to immobilized rGly m 4 with rGly m $4_{S111P/L150P}$ as inhibitor.

Binding of IgE antibodies of serum pools generated from 27 subjects with (group 1, positive pool – green) and 20 subjects without (group 2, negative pool – red) soybean allergy. IgE binding to immobilized rGly m 4 was inhibited by increasing concentrations of rGly m $4_{S111P/L150P}$ in ELISA. All curves were fitted by a 4-parameter, logistic sigmoidal curve fit.

rGly m $4_{S111P/L150P}$ showed EC_{50} values of 89.73 ± 11.69 ng for positive and 120.40 ± 17.28 ng for negative serum pool, respectively. Therefore with both serum pools a comparable inhibition curve with rGly m $4_{S111P/L150P}$ as inhibitor was observable, similar to the usage of rGly m 4 as inhibitor (see 3.7), indicating no differences in avidities regarding both analyzed serum pools. In contrast, a more potent inhibition of IgE binding was detected with rGly m 4 compared to rGly m $4_{S111P/L150P}$ (see Figure 27). Overall no potential importance of avidity comparing patients with and without soybean allergy was observable with both rGly m 4 and rGly m $4_{S111P/L150P}$ at least in this study population.

Next, IgE binding characteristics of both misfolded variants rBet v $1a_{S112P/R145P}$ and rGly m $4_{S111P/L150P}$ were addressed. For this, ten sera of patients with specific IgE against Bet v 1a or birch pollen were tested in combination with rBet v $1a_{S112P/R145P}$ variant. ImmunoCAP™ data of the respective patients are shown in Table 20.

Table 20: IgE levels determined by ImmunoCAP of ten subjects with sensitization to Bet v 1a or birch pollen.

Patient No.	sex	IgE (kU_A/L) determined by ImmunoCAP™		
		Total	Birch-specific (class)	sIgE Bet v 1 (class)
1	f	51	n.d.	21.5 (4)
2	m	206	n.d.	48.2 (4)
3	m	>5000	n.d.	32.2 (4)
4	m	>5000	n.d.	13.7 (3)
5	f	138	n.d.	67.3 (5)
6	f	825.0	>100 (6)	n.d.
7	m	858.0	>100 (6)	n.d.
8	m	498.0	>100 (6)	n.d.
9	f	699.0	>100 (6)	n.d.
10	f	852.0	>100 (6)	n.d.

Patient number, sex and levels of total serum IgE, birch- or Bet v 1-specific IgE in kU_A/L with resulting CAP class shown in brackets. All measurements were performed with a Phadia250 ImmunoCAP™. n.d. – not determined.

All patients showed a level of total IgE between 51 and >5000 kU_A/L with a median of 762 kU_A/L. For patients 1-5 level of Bet v 1-specific IgE was determined. These subjects are sensitized to Bet v 1 with an median of 32.2 kU_A/L and IgE levels between 13.7 and 67.3 kU_A/L. For remaining five subjects (patients 6-10) specific IgE level against birch pollen was determined with >100 kU_A/L (CAP class 6) classifying them as high responders to birch pollen. All patients (1-10) had a convincing history of pollinosis to early flowering tree pollen. At the time of this PhD thesis the detailed information on allergic symptoms of patients as well as information on allergic reactions to further allergens from food or pollen had not yet been revealed. Analysis of IgE binding with rBet v $1a_{S112P/R145P}$ and birch-sensitized patients in immunoblot is shown in Figure 32.

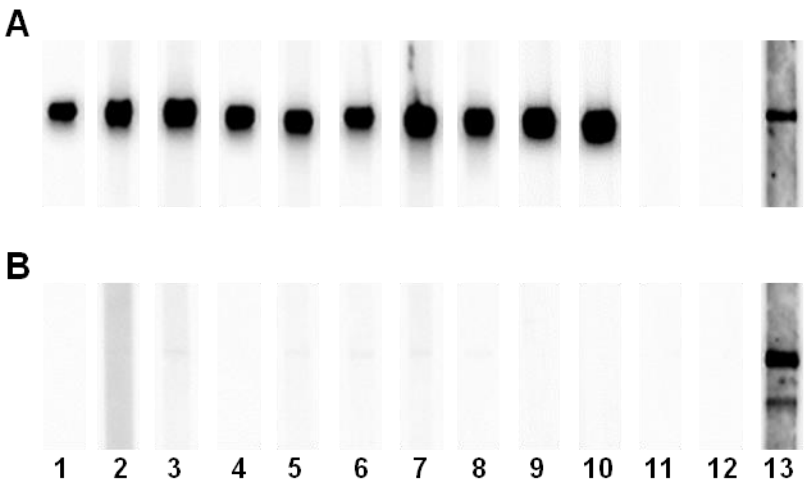


Figure 32: rBet v 1a but not rBet v 1a$_{S112P/R145P}$ show IgE binding with birch-allergic patients.

Ten sera of subjects with birch pollen allergy (lanes 1-10) analyzed in immunoblotting for IgE binding with rBet v 1a (A) and rBet v 1a$_{S112P/R145P}$ (B). Control experiments with a serum of a non-allergic patient (lane 11) and buffer control (lane 12) are shown as well as detection of Bet v 1-specific IgG antibodies (lane 13).

rBet v 1a but not rBet v 1a$_{S112P/R145P}$ was able to bind serum IgE from patients 1-10 in immunoblot analysis. Therefore native protein conformation was necessary for IgE binding at least with tested serum samples and misfolded protein conformation of rBet v 1a$_{S112P/R145P}$ might cause hindrance of interaction with IgE due to changes in existing IgE epitopes. As expected no IgE binding with both proteins was observable with serum of a non-allergic patient and buffer control with no serum used. Both rBet v 1a and rBet v 1a$_{S112P/R145P}$ were able to bind polyclonal Bet v 1-specific IgG in comparable intensity. This verifies the ability of rBet v 1a$_{S112P/R145P}$ to generally interact with immunoglobulin antibodies. Analysis with Bet v 1-specific IgG antibody revealed an additional signal below 18 kDa representing a degradation product of rBet v 1a$_{S112P/R145P}$. In addition to patients sensitized only to Bet v 1a and/or birch pollen allergen 36 subjects (groups 1 and 2), which were already analyzed with NCS and rGly m 4 proteins, were tested for IgE binding with rBet v 1a$_{S112P/R145P}$ and rGly m 4$_{S111P/L150P}$. In Figure 33 examples of nine patients are shown.

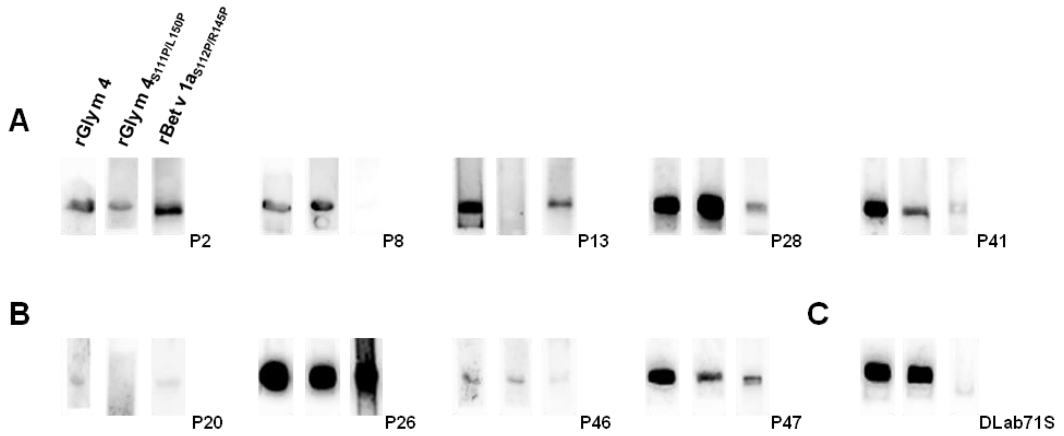


Figure 33: IgE binding of rGly m 4 and rBet v 1a variants.

Five representative patients sensitized and soy allergic (A) as well as four representative patients sensitized but without soybean allergy (B) and purchased DLab71S serum from a subject with birch and soy sensitization (C) were analyzed in immunoblot with rGly m 4, rGly m $4_{S111P/L150P}$ and rBet v $1a_{S112P/R145P}$. Number of each patient is indicated below each blot.

As already shown in 3.1 and 3.6.3 most patients of both groups showed IgE binding with rGly m 4. Different binding characteristics were observed in patients with and without soybean allergy concerning both misfolded variants. Some patients of group 1 (No. 2, 8, 9, 13, 33, 34 and 37) and group 2 (No. 20, 25, 45, 46 and 48) exhibited IgE binding with only rBet v $1a_{S112P/R145P}$ or rGly m $4_{S111P/L150P}$. In contrast both misfolded variants bound serum IgE in 59% (13/22) of subjects with and in 50% (7/14) of subjects without soy allergy. Only patients 7 and 40 from group 1 and patient 24 from group 2 had IgE binding only with rGly m 4 and neither with rBet v $1a_{S112P/R145P}$ nor rGly m $4_{S111P/L150P}$. Therefore both misfolded variants were still able to bind IgE antibodies but binding was dependent on subject's specific IgE antibody repertoire. Patient 44 from group 2 only had serum IgE against rBet v $1a_{S112P/R145P}$ but showed no IgE binding with rGly m 4 and rGly m $4_{S111P/L150P}$. Table 21 lists the results of immunoblots with sera of both groups of patients and both misfolded allergen variants. In addition a commercially available serum of a subject sensitized against Gly m 4 and Bet v 1 (DLab71S) was tested showing IgE interaction with rGly m 4 and rGly m $4_{S111P/L150P}$ but not with rBet v $1a_{S112P/R145P}$.

Table 21: Summary of IgE binding with misfolded variants of rGly m 4 and rBet v 1a.

	No.	rGly m 4	rGly m 4 S111P/L150P	rBet v 1a S112P/R145P	No.	rGly m 4	rGly m 4 S111P/L150P	rBet v 1a S112P/R145P
Group 1	1	X	X	X	22	X	X	X
	2	X		X	28	X	X	X
	3	X	X	X	30	X	X	X
	4	X	X	X	33	X	X	
	5	X	X	X	34	X	X	
	7	X			35	X	X	X
	8	X	X		36	X	X	X
	9	X	X		37	X	X	
	10	X	X	X	38	X	X	X
	11	X	X	X	40	X		
	13	X		X	41	X	X	X
Group 2	16	X	X	X	44			X
	20	X		X	45	X	X	
	24	X			46	X	X	
	25	X	X		47	X	X	X
	26	X	X	X	48	X	X	
	27	X	X	X	49	X	X	X
	43	X	X	X	50	x	X	X

Sera of patients with (22 subjects, group 1) and without (14 subjects, group 2) soybean allergy were analyzed in immunoblots with rGly m 4, rGly m $4_{S111P/L150P}$ and rBet v $1a_{S112P/R145P}$. Detected IgE binding is marked with (X).

3.8.4 Analysis of different molar ratios of misfolded variants

Because of different results obtained with folded and misfolded variants concerning IgE binding the impact of misfolded allergen on interaction with IgE was studied in more detail. For this, mixtures with defined molar ratios of rBet v 1a and rBet v $1a_{S112P/R145P}$ or rGly m 4 and rGly m $4_{S111P/L150P}$ were tested for their IgE binding characteristics. Results of physicochemical and immunobiological assays can then be transferred to purified allergen preparations used in diagnosis and therapy of

birch/soy allergy. During the purification process of a recombinant protein often an unknown ratio of misfolded protein might be generated in parallel not able to bind IgE compared to native-folded protein. Therefore, to ensure suitability and stability of allergen preparations it is important to verify the extent of misfolded protein and associated changes in IgE binding characteristics. For each rBet v 1a and rBet v $1a_{S112P/R145P}$ three protein preparations were used to generate three independent combinations between both allergens. For rGly m 4 and respective variant rGly m $4_{S111P/L150P}$ only one combination was used.

First, mixtures of folded and misfolded allergens of rBet v 1a and rGly m 4 were analyzed in CD spectroscopy. Ratios from 100% to 0.01% of misfolded variant were measured with a total number of ten and 14 molar ratios of rBet v 1a and rGly m 4, respectively. In Figure 34 CD spectra of rBet v 1a (A) and rGly m 4 (B) as well as changes at characteristic wave lengths are shown.

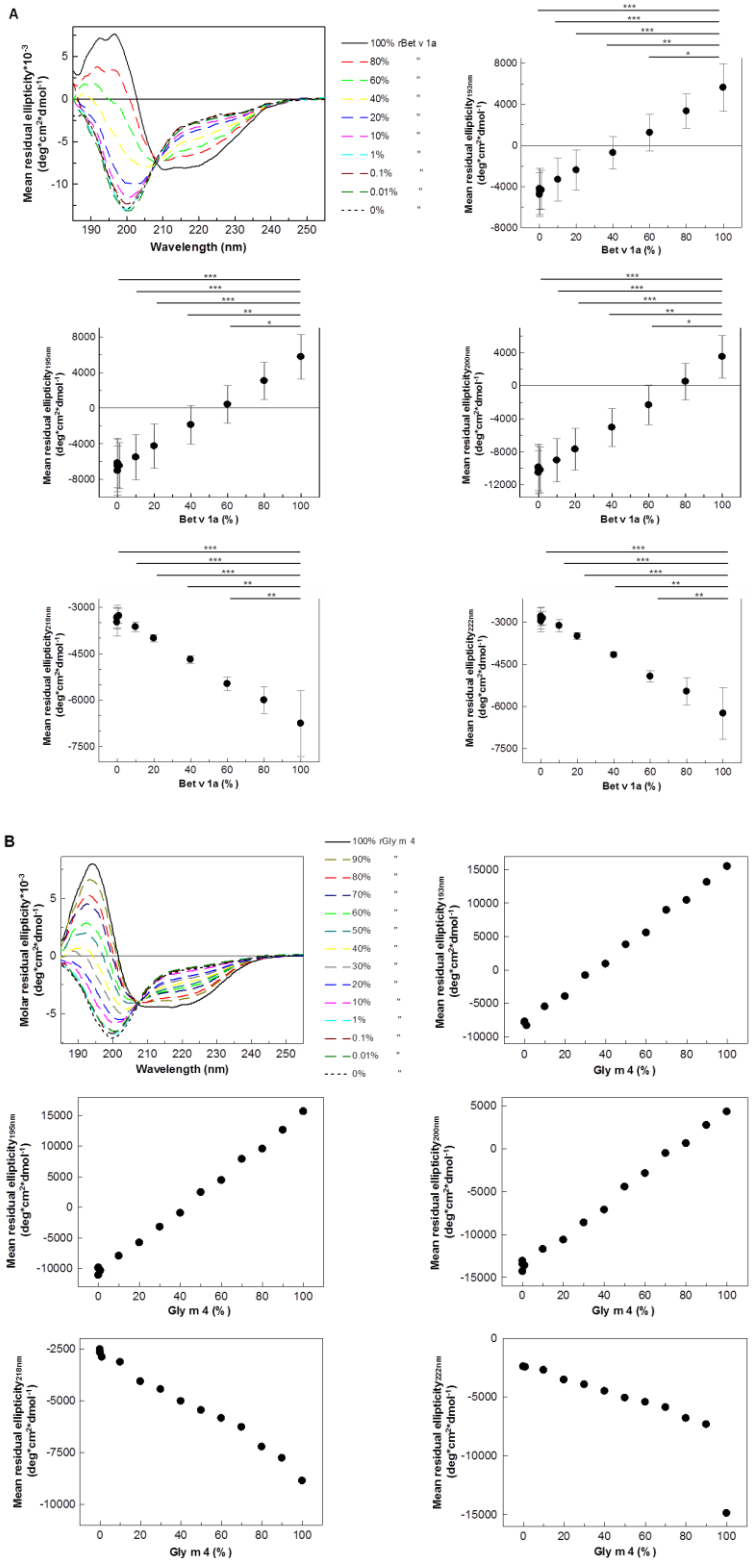


Figure 34: Circular dichroism of different molar ratios of rBet v 1a and rGly m 4 with their respective misfolded variants.

Spectra of combinations between rBet v 1a and rBet v 1a$_{S112P/R145P}$ (A) as well as rGly m 4 and rGly m 4$_{S111P/L150P}$ (B) are shown. For rBet v 1a ten and for rGly m 4 14 combinations from 0% to 100% rBet v 1a$_{S112P/R145P}$ or rGly m 4$_{S111P/L150P}$ were measured, respectively. Mean residual ellipticities at 193 nm, 195 nm, 200 nm, 218 nm and 222 nm are indicated for rBet v 1a and rGly m 4 with 95% confidence interval and statistical evaluations for measurements with rBet v 1a which were performed in three individual allergen combinations.

With increasing ratio of misfolded variant, spectra of both allergens changed from Bet v 1-type to one characteristic for unstructured proteins. Mixtures of rBet v1a/rBet v $1a_{S112P/R145P}$ and rGly m 4/rGly m $4_{S111P/L150P}$ both showed an isodichroic point at about 208 nm. For detailed analysis specific wavelengths characteristic for secondary structure elements were examined. 193 nm and 222 nm with positive and negative bands in α-helical proteins as well as 195 and 218 nm characteristic for positive and negative ellipticities in β-sheets were analyzed (Holzwarth & Doty, 1965; Greenfield & Fasman, 1969). In addition wave length at 200 nm showing negative bands characteristic of both misfolded variants was used (Venyaminov *et al.*, 1993). An increase (193, 195, 200 nm) or decrease (218, 222 nm) with increasing ratio of well-folded allergen was observable in both allergen combinations. For rBet v 1a mixtures a comparison between actual and theoretical measurements was possible. Measured ellipticities at analyzed characteristic wave lengths correlated well with those of theoretical combinations, at least from 100% to 10% of rBet v 1a. Below 10% of rBet v 1a only low correlation of actual-theoretical comparison is observable. Statistical analysis of rBet v 1a combinations revealed that mixtures containing 60% of rBet v $1a_{S112P/R145P}$ could be resolved from 100% rBet v 1a at wave lengths 193, 195, 200 and 218 nm with 95% confidence interval. At 222 nm however only 40% of rBet v $1a_{S112P/R145P}$ was well-separated from 100% rBet v 1a. Due to analysis of only one combination between rGly m 4 and rGly m $4_{S111P/L150P}$ no comparison of actual/theoretical values and statistical evaluation was possible.

Combinations of both rBet v 1a/rBet v $1a_{S112P/R145P}$ and rGly m 4/rGly m $4_{S111P/L150P}$ were further analyzed for IgE binding in immunoblots. For rBet v 1a a pool of ten subjects (see Table 20) was used while DLab71S serum (see Figure 33) was applied for rGly m 4. Results of both immunochemical analyses are shown in Figure 35.

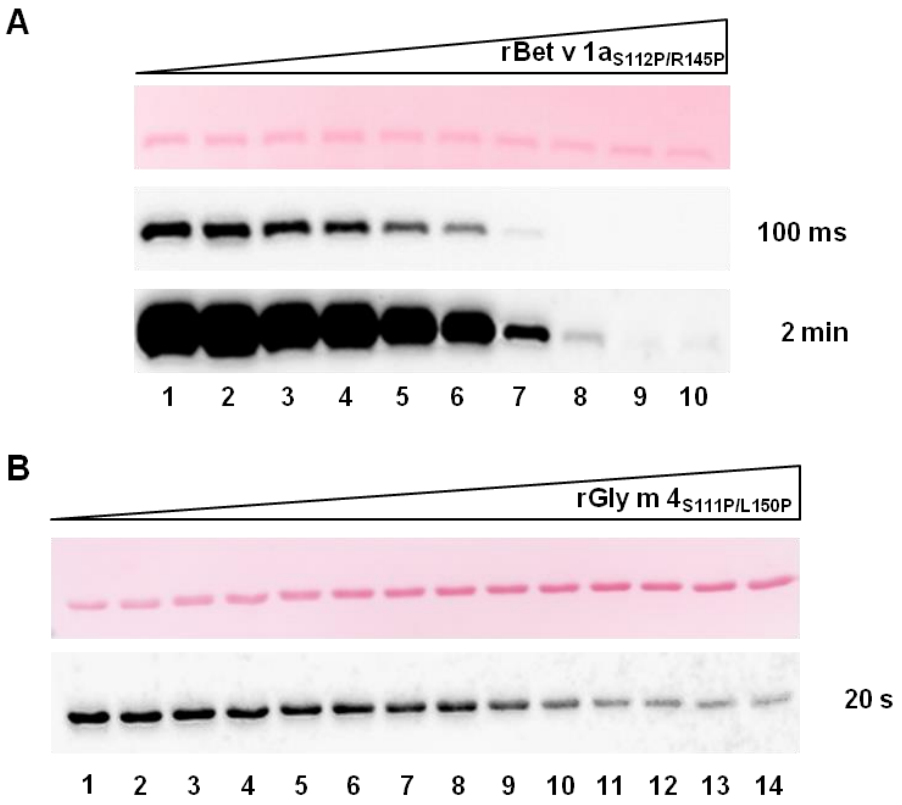


Figure 35: Increasing ratios of misfolded variants show a decrease in IgE binding in both rBet v 1a and rGly m 4 combinations.

Immunoblot of ten mixtures of rBet v 1a/rBet v 1a$_{S112P/R145P}$ (A) and 14 combinations between rGly m 4/rGly m 4$_{S111P/L150P}$ (B) were analyzed with pool serum of ten patients and DLab71S serum, respectively. Increasing ratios of each misfolded variant were transferred and Ponceau S stained. Chemiluminescence after 100 ms and 2 min for rBet v 1a and 20 s for rGly m 4 are shown.

With increasing ratio of misfolded variants a decrease of IgE binding is observable in both rBet v 1a and rGly m 4 mixtures, as expected. However, a difference in IgE signals was found at 100% of rBet v 1a$_{S112P/R145P}$ and rGly m 4$_{S111P/L150P}$, respectively. Concerning the pool of sera no IgE binding could be detected at 100% of rBet v 1a$_{S112P/R145P}$ confirming results of immunoblot with each individual serum alone. In contrast 100% rGly m 4$_{S111P/L150P}$ showed residual IgE binding. This also verifies the results obtained earlier (see 3.8.3) where DLab71S serum showed IgE binding with both rGly m 4 and rGly m 4$_{S111P/L150P}$. Comparison of actual and theoretical data for rBet v 1a mixtures revealed an over-interpreted signal of IgE binding especially at ratios of rBet v 1a ≤ 1%.

IgE binding of different molar ratios of rBet v 1a and rBet v 1a$_{S112P/R145P}$ was furthermore analyzed in ELISA experiments shown in Figure 36.

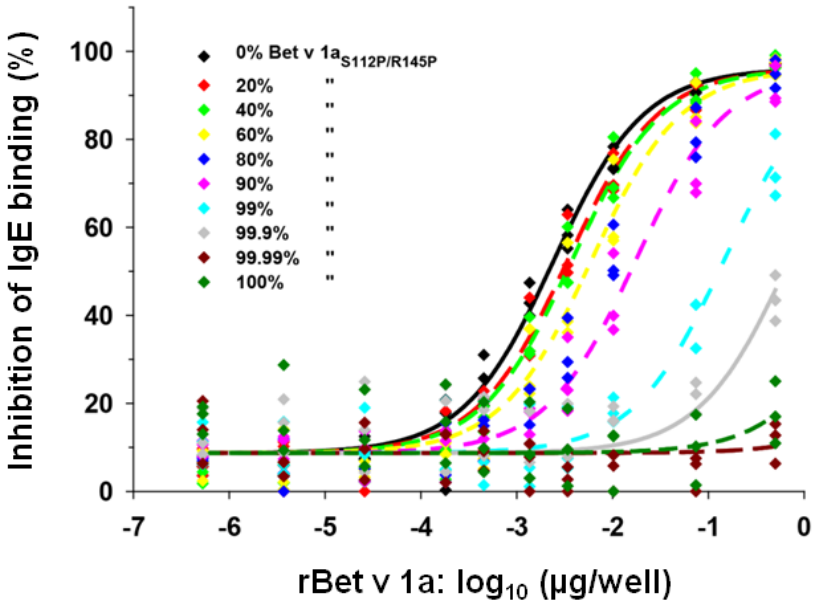


Figure 36: Increasing ratio of rBet v 1a$_{S112P/R145P}$ causes a decrease in inhibition of IgE binding.

Binding of IgE antibodies from a serum pool of ten subjects (see Table 20) to immobilized rBet v 1a was inhibited by different concentrations of mixtures between rBet v 1a and rBet v 1a$_{S112P/R145P}$ in ELISA. All curves were fitted by a 4-paramter, logistic sigmoidal curve fit with same slope as well as same lower and upper asymptote.

As expected from previous results in immunoblot experiments rBet v 1a but not rBet v 1a$_{S112P/R145P}$ is able to inhibit IgE binding with rBet v 1a. Combinations of both allergens showed a concentration-dependent decrease in inhibition of IgE binding verified by a shift in half-maximal inhibition (EC_{50}). Each curve was fitted with a 4-parameter logistic model assuming sigmoidal curve fit with same lower and higher asymptote. Estimated common slope was 0.91 with a 95% confidence interval of 0.85-0.96. Comparison between actual an theoretical EC_{50} values were between 1.09 at 20% rBet v 1a$_{S112P/R145P}$ and 0.29 at 99.9% rBet v 1a$_{S112P/R145P}$. No correlation could be determined at concentrations of 99.99% and 100% rBet v 1a$_{S112P/R145P}$ due to non-evaluable inhibition curves.

Next, a cellular system using humanized rat basophil leukaemia (RBL) cells was used to determine the biological activity of different molar ratios of rBet v 1a and rBet v 1a$_{S112P/R145P}$. With RBL cell release assay the release of mediator molecules of cells sensitized with serum IgE can be measured in correlation to added allergen. Resulting release curves are shown in Figure 37.

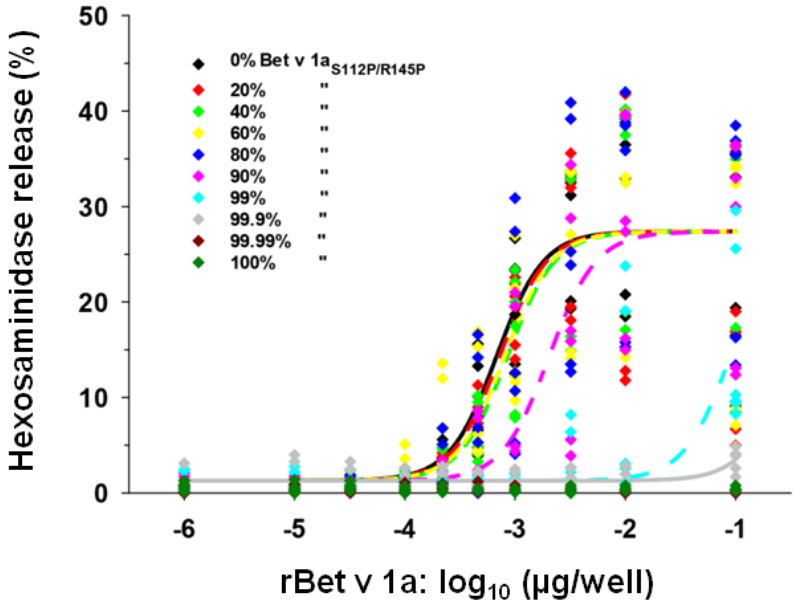


Figure 37: Increasing concentrations of rBet v 1a$_{S112P/R145P}$ cause a reduction in release of mediator molecules in rat basophil leukaemia cells.

Humanized rat basophil leukaemia cells were sensitized with the pool of ten subjects showing IgE against rBet v 1a and/or birch pollen. Release of β-hexosaminidase molecules was determined dependent on concentration of different ratios between rBet v 1a and rBet v 1a$_{S112P/R145P}$ causing a cross-linking of membrane-bound IgE molecules.

Again EC_{50} values for each combination of rBet v 1a and rBet v 1a$_{S112P/R145P}$ were measured and compared with theoretical values. Overall, a release of mediator molecules was observable at concentrations of 100% to 1% rBet v 1a. Hereby actual-theoretical values ranged from 0.86 to 1.32 (80-1% rBet v 1a). No reliable release could be determined at ratios of 0.1% rBet v 1a and below. No significant differences could be observed between the EC_{50} values for 100% rBet v 1a and all rBet v 1a/rBet v 1a$_{S112P/R145P}$ compositional protein preparations. According to a confidence interval of 95% only the rBet v 1a/rBet v 1a$_{S112P/R145P}$ protein combination containing 1% rBet v 1a resolved analytically from the mediator release obtained with 100% rBet v 1a.

Finally, actual/theoretical values of CD spectroscopy, immunoblot, ELISA and RBL cell release assay were compared for rBet v 1a/rBet v 1a$_{S112P/R145P}$ ratios between 10-100% rBet v 1a (Figure 38).

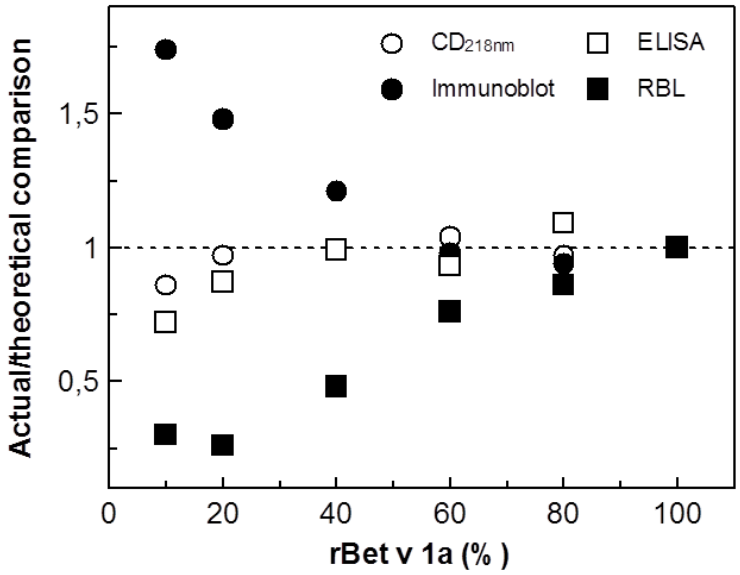


Figure 38: Summary of actual/theoretical comparisons in CD, immunoblot, ELISA and RBL cell release.

Values of actual/theoretical measurements concerning different molar ratios between rBet v 1a and rBet v 1a$_{S112P/R145P}$ from 10% to 100% rBet v 1a are shown. Dotted line indicates identical actual/theoretical values.

Values for actual/theoretical comparison were comparable between CD spectroscopy at 218 nm and ELISA with underestimation especially at concentrations ≤20% rBet v 1a. In contrast, results of immunoblot analysis showed overestimated values at ratios of ≤40% rBet v 1a but a good actual/theoretical correlation at high amounts of rBet v 1a. Weakest correlation between actual and theoretical values was found in RBL cell release assay with underestimation increasing from 80% to 10% of rBet v 1a. Overall, a large variation especially in measurements with concentrations below 10% of rBet v 1a was observable. Total values are 0.86 for CD spectroscopy at a wavelength of 218 nm, 1.74 for immunoblotting, 0.72 for ELISA and 0.3 for RBL cell release assay. Values for all actual-theoretical comparisons for each assay are shown in Table A 4 in the appendix.

4 DISCUSSION

4.1 Correlation of serum antibody levels and clinical phenotype

After thorough anamnesis and skin prick tests, measurement of allergen-specific serum IgE level is often the first immunological parameter to be determined in allergy diagnosis. Patient's serum is analyzed with a standardized assay like ImmunoCAP™ or any related test system. Often the resulting quantitative allergen-specific IgE levels are correlated directly with the respective allergy and even with the severity of allergic symptoms (Sampson & Ho, 1997; Sampson, 2001). However, a direct link between IgE level and clinically confirmed allergy remains controversial (Asero *et al.*, 2007; Guhsl *et al.*, 2015). Recently it has been shown that IgE levels specific for Bet v 1-related allergen Mal d 1 which causes birch-related allergy to apples do not correlate with allergic symptoms (Ebo *et al.*, 2010). Nevertheless a proper diagnosis of the allergy is indispensable to avoid putative unnecessary medication or restrictions in the patient's life. Furthermore the level of allergen-specific IgG_4 antibodies in sera of patients receiving SIT is often accompanied with a reduction of allergic symptoms due to blocking of IgE-mediated allergic symptoms (Tomicić *et al.*, 2009).

In this study the IgE and total IgG levels specific for Bet v 1, Gly m 4, Gly m 5 and Gly m 6 in sera of patients with and without clinically confirmed soybean allergy were determined to analyze a possible correlation between serum antibody levels and allergy (see 3.1). A total number of 47 patients sensitized to Gly m 4 and allergic to birch pollen were included in the study. 27 subjects (group 1) showed mild to severe symptoms after soybean consumption and 20 patients (group 2) were sensitized to Gly m 4 but not allergic to soy, as confirmed by provocation testing.

In ImmunoCAP™ analysis two patients (No. 34 and 48) did not have sIgE to Bet v 1 or Gly m 4 despite a diagnosed birch pollen allergy. For patient 34 a true relationship between birch and soy allergy was already questioned (Kleine-Tebbe *et al.*, 2002). In addition patient 48 had no IgE against Gly m 4 or Bet v 1 but had been clinically characterized as allergic to soy. ImmunoCAP™ data of residual 45 patients were characteristic for subjects with birch pollen allergy showing sIgE levels for the major allergen of birch Bet v 1 ≥CAP class 3. Furthermore these patients had sIgE against soybean allergen Gly m 4. Therefore, irrespective of clinical symptoms after soy consumption, majority of patients were sensitized to both allergens. This is a common phenomenon as allergen-specific IgE *per se* does not indicate clinical

relevance of a sensitization (Guhsl *et al.*, 2015). Nevertheless some controversial studies about a potential correlation between severity of symptoms and allergen-specific IgE levels are discussed. In contrast to birch-related allergy to apple where no correlation between specific IgE levels and the severity of symptoms was found, allergic reactions in peanut allergy correlate well with specific IgE level (Sampson & Ho, 1997; Sampson, 2001; Asero *et al.*, 2007).

Only one patient of the study group had low levels of sIgE directed against Gly m 5 and Gly m 6 but high levels of Bet v 1- and Gly m 4-specific IgE, suggesting a birch-related soy allergy rather than a class I soybean allergy. Therefore symptoms after DBPCFC with soybean can be linked to a pollen-related soybean allergy rather than a classical allergy to soy. Comparing the median levels of sIgE, patients of group 1 had higher levels of sIgE against both Bet v 1 (37.0 kU_A/L) and Gly m 4 (9.3 kU_A/L) compared to 29.6 kU_A/L for Bet v 1 and 4.5 kU_A/L for Gly m 4 in subjects merely sensitized but without clinical soy allergy. In contrast, in both groups sIgE levels ranging from 0 to >100 kU_A/L for Bet v 1 and Gly m 4 were observed. Concerning presence and level of sIgE directed against Bet v 1 and Gly m 4 no difference in both groups of patients was observable. These observations confirmed that a birch-related soybean allergy cannot be diagnosed based on allergen-specific IgE against Gly m 4 only. Furthermore 70% of subjects from group 1 had severe symptoms like dyspnea, flush, face swelling, throat tightness or urticaria after soybean consumption but sIgE levels did not correlate with severity of allergic symptoms. Patients with high levels of sIgE against Gly m 4 (46.3 kU_A/L, patient 2) had mild symptoms like OAS, whereas subjects with a relatively low amount of Gly m 4-specific IgE (4.0 kU_A/L, patient 33) exhibited severe reactions like face swelling and throat tightness. Similar results were shown in birch-related apple allergy, were no correlation between sIgE and the severity of allergic symptoms after apple consumption was observed (Ebo *et al.*, 2010). Until today the results of several studies performed to differentiate between patients with and without birch pollen-allergic food allergy by measurement of allergen-specific IgE remain controversial (Reuter *et al.*, 2006; Asero *et al.*, 2007; Hansen *et al.*, 2009; Ebo *et al.*, 2010; Ballmer-Weber *et al.*, 2012; Tolkki *et al.*, 2013; Guhsl *et al.*, 2015).

Furthermore IgE immunoblotting was performed as basic and serum-saving method to screen all patients of the study population in a short period of time. Major drawback of immunoblot compared to ELISA experiments is the use of denatured

allergens but for IgE binding to rGly m 4 a native protein conformation is required and therefore only properly refolded allergens are detected via immunoblot. Results obtained by ImmunoCAP™ were in line with allergen profiles generated by immunoblots. Again there was no correlation between the presence of birch-related soy allergy and Gly m 4-specific IgE antibodies. This was already shown by Ballmer-Weber *et al.* where no correlation between symptoms and IgE immunoblotting results was detected in birch-soy allergy (Ballmer-Weber *et al.*, 2007). In total 83% (30/36) of patients from groups 1 and 2 showed IgE binding to rGly m 4 in immunoblot. Residual six patients exhibited low Gly m 4-specific IgE levels according to ImmunoCAP™ with no (subjects 25, 45, 46, 48 and 49) or mild (patient 34) symptoms upon consumption of soy. It is conceivable that the Gly m 4-specific IgE levels of the sera of these patients were below the detection limit of the immunoblot system used. These results indicate the potential of a more sensitive analysis of IgE levels with ImmunoCAP™ compared to immunoblot using recombinant allergens which was already demonstrated by Mittag *et al.* (Mittag *et al.*, 2004b). Nevertheless, presence of IgE antibodies directed against Gly m 4 in both assays does not necessarily refer to a birch-related soybean allergy. Thus the evaluation of Gly m 4-specific IgE levels *per se* is not recommended as diagnostic tool for birch-related soybean allergy. However, quantitative rGly m 4-specific IgE signals obtained by immunoblot analysis correspond with the IgE levels determined in ImmunoCAP™. In immunoblots a complex pattern of different IgE-binding proteins was observed when analyzing soy extract. These signals refer to at least 18 different known IgE-binding proteins present in soybean extract. To date only Gly m 1 to Gly m 8 are accepted as allergens by IUIS. Patients might have IgE antibodies against multimeric soy seed storage proteins Gly m 5 or Gly m 6 which have a monomeric molecular weight in the range of 50-70 kDa and could be measured at least in small amounts (between 0 and 0.35 kU_A/l) in ImmunoCAP™. Furthermore presence of antibodies directed against cross-reactive carbohydrate determinants (CCD), carbohydrate moieties of allergenic molecules, which might be able to induce an allergic response, are possible. Availability of such antibodies might be verified with western blot inhibition experiments or specific ImmunoCAP™ measurements. Furthermore ImmunoCAP™ performed with soybean extract might be used, for example with patients 5, 40, 41 and 43, to measure IgE binding to other proteins than only Gly m 4. In total, IgE binding patterns to soy proteins of subjects from both groups are very

individual and complex without any correlation to the severity of allergic reactions. Again several patients with sensitization to Gly m 4, as determined by ImmunoCAP™, did not show an IgE signal representing Gly m 4 in immunoblot using soy extract. This might be caused by the low amount of Gly m 4 protein present in soybean extract as already determined by Mittag *et al.* using soybean ImmunoCAP™ analyses (Mittag *et al.*, 2004b).

Both groups of patients had comparable levels of total Gly m 4-specific IgG as determined in ImmunoCAP™. Neither a correlation of sIgE levels nor Gly m 4-specific IgG levels and clinical reactivity to soybean was observed. Highest level of IgG directed against Gly m 4 was found in a patient of group 2 (patient 16) without symptoms to soy. This is in line with findings that IgG antibodies might act as blocking antibodies inhibiting allergen-IgE binding causing no/mild allergic symptoms (James *et al.*, 2011; Jutel *et al.*, 2013). Patients with mild symptoms like OAS might have high or low Gly m 4-specific IgG levels of 36.9 mg_A/L (patient 11) and 7.3 mg_A/L (patient 2). These Gly m 4-specific IgG levels were also observed in sera of patients exhibiting severe reactions to soy. Patient 37 with 35.8 mg_A/L of Gly m 4-specific IgG exhibited dyspnea while patient 28 with only 8.0 mg_A/L sIgG showed face swelling after soy consumption. For a thorough allergen profile, measurement of IgG_4 levels for both Bet v 1 and Gly m 4 is necessary. Recently it was shown that ratios between IgG_4 and IgE in Bet v 1-related food allergens correlate with tolerance to several foods (Geroldinger-Simic *et al.*, 2011). Therefore analysis of correlation between IgG_4 level or IgG_4/IgE ratio and allergic symptoms might be interesting in birch-related soy allergy. Nevertheless influence of IgG_4 in outbreak and severity of allergies remains controversial as in 2015, Guhsl *et al.* reported a comparable level of IgG_4 and IgG_4/IgE ratio in patients with and without soybean allergy while an increase in $sIgG_4$ levels during SIT is often observed in successful immunotherapy (Guhsl *et al.*, 2015).

4.2 Generation of a preliminary epitope profile of Gly m 4

For the identification of IgE epitopes a variety of different approaches and methods have been used to date. Briefly, these methods differ mainly in their effort, bioinformatical content, costs and output data. In this study, a combination of experimental and bioinformatically-based methods were used to identify putative IgE epitopes of Gly m 4.

4.2.1 Identification of candidate residues involved in IgE binding

Before distinct IgE epitope profiles can be determined properly, candidate residues which might dominate the energetic of allergen-antibody binding (functional epitope) need to be identified to generate a preliminary epitope profile (Dall'Acqua *et al.*, 1998). This concept was already proven by changing few surface-exposed residues in a non-IgE-binding protein with structural homology to Bet v 1 to a cross-reactive binding site for IgE (Berkner *et al.*, 2014). Such residues may be located all over Gly m 4 amino acid sequence making a thorough analysis of Gly m 4 necessary. This was carried out by a combination of experimental and theoretical approaches (see 3.2).

Phage-display technology is an experimental-based method to identify antibody-binding peptides by usage of a phage-displayed peptide library. Phage-display was used successfully with several allergens like Ara h 1, Ara h 8, Bet v 1, Gly m 4 or Pru av 1 (Mittag *et al.*, 2006; Bøgh *et al.*, 2014). Based on an established protocol, described by Mittag *et al.*, phage-display and competitive immunoscreening with rGly m 4 was performed (see 3.2.1) (Mittag *et al.*, 2006). However analysis of the data always revealed a minimum of 20% of false-positives, which are eluted phages without any presented peptide. Such phages are present with a ratio of 10% in commercially available library, as stated by manufacturer, but should not be eluted by Gly m 4-specific elution because an inserted peptide is a prerequisite for binding of phages to the IgE antibody molecule. For an increase in specificity of eluted phages several modifications were included in the established protocol. But neither an increased number of washing steps nor the usage of different washing and blocking buffers containing additional Tween20 and/or BSA reduced number of unspecifically eluted phages. Furthermore additional controls were performed by amplification and growth on IPTG/Xgal-plates to check for the presence of unspecifically eluted phages. Finally, negative selection process, where amplified phages after competitive elution with rGly m 4 were incubated with magnetic beads and anti-human IgE-coupled beads to remove unspecific phages was introduced. Despite these modifications in phage-display protocol the total number of false-positives could not be reduced. Furthermore false-positives were also observed when buffer or a non-atopic serum control instead of a serum containing Gly m 4-specific IgE antibodies was used. This indicates an unspecific interaction of phages with any of the materials used in experimental setup. Possible unspecific binding of phages might occur with

magnetic beads or anti-IgE antibodies. According to this, the phage-display approach was no longer pursued. Nevertheless published Gly m 4-based mimotope data from Mittag *et al.* were used for the identification of candidate residues (Mittag *et al.*, 2006).

These mimotopes were compared with results obtained by a theoretical bioinformatical approach where overlapping 7mer peptides representing total amino acid sequence of Gly m 4 were mapped onto allergen's molecular surface (see 3.2.1). Mimotopes are peptides that bind antibodies and thus mimic at least in part a conformational epitope of an antigen. In both approaches peptides were used, either identified experimentally via phage-display or based on amino acid sequence, as input data followed by usage of EpiSearch algorithm to map the peptide sequence onto the molecular surface of the allergen (Negi & Braun, 2009). Usage of linear peptides is currently performed in epitope analysis for identification of linear IgE-binding sites (Reese *et al.*, 2001; Shreffler *et al.*, 2005; Vereda *et al.*, 2010). Nevertheless even conformational IgE epitopes often show stretches of amino acids which are adjacent in primary sequence of the allergen. Therefore mapping of linear peptides might lead to the identification of putative conformational epitopes as well. This is in line with the mechanism of allergic response where the allergen is digested during gastrointestinal uptake into small peptides which are then presented e. g. on the surface of dendritic cells interacting with T cells resulting finally in an allergic response. Both approaches resulted in largely overlapping putative epitopes with center residues 29, 30, 37, 94, 119 and 150 identified most frequently and with a total number of 8-17 amino acids of each predicted epitope. For a further analysis of Gly m 4-specific IgE epitopes single residues rather than mimotopes were used to analyze specific amino acids potentially involved in IgE binding. Such functional epitopes dominate the energetics of allergen-IgE binding and substitution of a single residue might have a tremendous effect on interaction with antibody. Together a total number of 69 amino acids with potential impact on IgE binding was identified as candidate residues using both approaches.

In an additional analysis all residues of Gly m 4 were considered for potential IgE-binding ability theoretically (see 3.2.2). This was done using several inclusion criteria to choose only residues with potential relevance for IgE binding. Together with 69 candidate residues identified by experimental/theoretical screening as well as already published data of potential IgE-binding residues of Bet v 1 a resulting number

of 51 amino acids were chosen. Most of identified candidate residues showed a side chain with either charged or functional groups enabling for potential molecular interactions. Therefore especially aspartate, glutamate but also lysines were chosen. Lysine with its long side chain might act as an anchor by making the first contact with IgE antibodies. Furthermore amino acids with hydroxyl-, amino-groups or aromatic rings located on side chains were identified as putative IgE-binding residues. Candidate residues were distributed over the complete protein structure. The mechanism of effector cell activation in type I allergy requires at least two distinct IgE epitopes being present on Gly m 4 allowing simultaneous binding of two different IgE molecules for cross-linkage of the FCεRI receptor via allergen. Therefore the complete protein surface of Gly m 4 was considered to identify all possible putative IgE-binding sites. Within both approaches amino acids might be identified as candidate residues e. g. because of a characteristic side chain with potential importance for antibody binding, but with only a minor role in actual allergen-IgE interaction. Nevertheless all candidate residues need to be considered to minimize the risk of missing any putative binding site. False-positive candidate residues can be identified and removed by the next step in this analysis where each single amino acid is checked individually for its potential IgE binding in the context of a properly folded protein structure.

4.2.2 Generation of a NCS-based protein library

In this study, epitope grafting was used to test whether specific amino acids of Gly m 4 are critical for binding of IgE antibodies. Epitope grafting in the context of this study describes the transfer of putative or known antibody-binding sites from Gly m 4 to the Gly m 4-type non-allergenic recombinant derivative of norcoclaurine synthase (NCS) from meadow rue (see 3.3). NCS exhibits no/low binding of Bet v 1-specific serum IgE while being structurally homologous to Bet v 1 and Gly m 4 (Berkner *et al.*, 2014). If NCS variants presenting Gly m 4-specific amino acids bind IgE, the grafted residues are most likely critical for the interaction with IgE antibodies. Here I used a C- and N-terminally truncated variant of native NCS (Δ51NCS$_{N42/P49}$, Nürnberg, 2014, Master thesis) with an internal deletion of 22 amino acids and an insertion and substitution of two amino acids to ensure a PR10 protein-type length of 160 amino acids and secondary/tertiary structure. To date, the only known IgE

epitope of Gly m 4 was identified by epitope grafting using a variant of the NCS model protein that presented an epitope which is recognized by the Bet v 1-specific murine monoclonal antibody BV16 (Mirza *et al.*, 2000; Berkner *et al.*, 2014). Therefore, only individual candidate residues of Gly m 4 were grafted onto NCS. In this experimental setup individual amino acids of Gly m 4 which dominate the energetics of allergen-IgE binding within a structural epitope (functional epitopes) can be identified and analyzed. Thus, these grafted amino acids might mimic an antibody-binding site of Gly m 4 inducing specific binding of serum IgE with model protein NCS. Nevertheless these single residues do not bind an IgE antibody alone but might adapt the surrounding molecular surface of NCS to a more Gly m 4-type binding site enabling the interaction of NCS variants with IgE antibodies. Epitope grafting onto a non-antibody binding allergen-type model protein is superior to classical mutational epitope analysis of the allergenic protein, as it circumvents a polyclonal IgE response which might mask the impact of individual epitopes with low antibody binding frequency. For epitope grafting a detailed structural overlay of both proteins, NCS and Gly m 4, was required to identify structurally homologous residues that might be suitable as candidates for a comprehensive epitope analysis of Gly m 4.

4.3 Screening of NCS library

To identify functional IgE epitopes, variants from a NCS protein library were used for a systematic and comprehensive epitope analysis (see 3.4). With this initial screening, it is possible to identify NCS variants with IgE-binding characteristics which can be characterized in a more detailed analysis using purified NCS variants afterwards. In this study, screening of non-purified inclusion bodies of NCS variants was performed in immunoblot, suitable for screening of 48 variants together with sera of 47 patients with (27 subjects) and without (20 subjects) soybean allergy (see 3.4.2). Assuming that recombinant NCS variants show correct refolding and therefore Bet v 1-type protein conformation, observed IgE binding directly correlates with the substitutions of Gly m 4-specific amino acids. Whenever a single residue within an epitope dominates the energetic of binding to serum IgE, this particular amino acid might act as a functional epitope for IgE. In Dall'Acqua *et al.* this concept of functional epitopes was described the other way round, where a structural epitope comprising 12 amino acids of lysozyme interacting with an antibody fragment was

determined and with a thorough analysis using single variants of lysozyme showing individual substitutions in this structural epitope, two residues, namely Q121 and R125, could be identified as those amino acids energetically critical for antibody binding (Dall'Acqua *et al.*, 1998). This observation justifies the use of Gly m 4-specific NCS variants to identify and analyze functional epitopes of Gly m 4. Major drawback of immunoblot compared to ELISA experiments is the use of denatured allergens but for IgE binding to rGly m 4 a native protein conformation is required and therefore only properly refolded allergens are detected via immunoblot

4.3.1 Characterization of identified potential functional IgE epitopes

In total 46 putative functional IgE epitopes were identified via immunoblot analysis (see 3.4.2). Using immunoblot together with inclusion bodies of expressed NCS variants a reliable and time-saving initial screening of all sera/NCS variants combinations was possible. As IgE binding requires native protein conformation only NCS variants which were able to properly refold into Bet v 1-type conformation upon SDS-PAGE and subsequent transfer to nitrocellulose membrane were detected. Consequently some differences in serum IgE binding due to an increased variability of blotted allergens or difficult evaluation of positive/negative signals in immunoblot of NCS variants as well as NCS model protein might occur but usage of immunoblot as initial experiment is justified because identified putative functional epitopes were critically analyzed via purified NCS variants using dot blot (see 3.5.2) and ELISA (see 3.5.3) afterwards. In addition inhibition experiments might be useful to verify Gly m 4 specificities of IgE signals, as single amino acid substitutions in NCS theoretically could result in binding of IgE antibodies directed against further allergens present in patient's serum (e. g. allergens directed against house dust mite). Taken together immunoblot analysis represents the most straight forward and crude approach considering a careful interpretation of observed IgE signals and a cautious analysis of obtained results via control experiments. Overall, selection of candidate residues by mapping of mimotopes and rational mutagenesis represented a suitable approach for identification of putative IgE-binding functional epitopes. Only grafted residues E95/K96 and D140 did not induce IgE binding in any of the patients' sera analyzed. 50% of identified putative functional IgE epitopes are Gly m 4-specific amino acids that are not found in Bet v 1a. Gly m 4-specific residues

are distributed over the molecular complete surface of Gly m 4. For initial contact between IgE and allergen, Gly m 4-specific residues but also residues identical in Gly m 4 and Bet v 1a can act as anchor to initiate IgE binding. Nevertheless distribution of Gly m 4-specific functional IgE epitopes indicates that IgE epitope patches might not be composed only by Gly m 4-specific residues. IgE interaction sites on Gly m 4 are more likely composed of both, Gly m 4-specific and Bet v 1a/Gly m 4-homologous residues. This is in line with the course of the allergic disease, where sensitization to Gly m 4 does not occur in patients without prior birch pollen sensitization. Patients with soybean allergy showed an average of 21 compared to 18 putative functional epitopes identified for subjects without allergic symptoms to soy. In addition to a slightly increased average number of potential functional IgE epitopes, patients with clinically confirmed soybean allergy also presented an increased range in putative functional epitopes. This is 12-32 compared to 14-26 putative functional epitopes in patients without soy allergy. From this a potential correlation between number of functional IgE epitopes and clinical reactivity might exist although differences between both groups are relatively small. Only three putative functional epitopes were identified individually with one group of patients. Grafted residues D92/T93 and E147 showed IgE binding solely with three patients having soybean allergy. In contrast N108 only gave IgE signal together with one serum of subjects sensitized but not soy allergic. Nevertheless relevance of these group-specific residues needs to be further evaluated. Both, D92/T93 and E147 are conserved in Bet v 1a, while N108 is only present in Gly m 4. D92 and T93 were already identified as parts of IgE-binding Bet v 1a peptides (Gieras *et al.*, 2011). E147 belongs to a putative epitope of Bet v 1a interacting with a monoclonal antibody (Jahn-Schmid *et al.*, 2005; Hecker *et al.*, 2012). Furthermore D92, T93 and E147 belong to different mimotopes identified by Mittag *et al.* (Mittag *et al.*, 2006). N108 was reported by Ganglberger *et al.* as part of an IgE epitope of Bet v 1 identified using mimotopes from a phage library (Ganglberger *et al.*, 2001). Besides this, no characteristics for the two groups of patients could be found concerning occurrence of potential functional epitopes. Rather, each patient showed an individual pattern of identified putative functional IgE epitopes.

Regarding subjects with soybean allergy no correlation between any putative functional epitopes or their total number and clinical symptoms could be found. Patients 9 and 11, both exhibiting OAS after soybean consumption differed markedly

regarding total number of IgE signals, with 26 and 14 different potential identified functional IgE epitopes, respectively. Subjects with severe symptoms like dyspnea and urticaria or dyspnea and face swelling as observed for patients 28 and 35, respectively, showed large differences with 11 compared to 20 identified putative functional epitopes in total. Overall only small differences in IgE binding frequencies regarding all patients as well as regarding subjects from group 1 or 2 were found. Therefore, analysis of putative functional epitope pattern did not lead to any conclusions concerning presence or severity of birch-related soy allergy. Rather a patient- or symptom-specific distribution of IgE epitopes is more likely.

NCS variants $\Delta 51NCS_{F3/N42/P49}$ and $\Delta 51NCS_{F5/N42/P49/D100}$ showed IgE binding with all used sera at least within initial screen using immunoblot. F3 is located at the N-terminus of primary sequence in first β-sheet and was introduced for a lysine in NCS. Phenylalanine is an aromatic and hydrophobic amino acid without any functional groups contributing especially to secondary structure stability via intramolecular interaction with lysine or arginine residues. These bindings arise from cation π-interactions between negative electrostatic potential of aromatic ring and cations like positively charged amino acids. Nevertheless phenylalanine can contribute to ligand binding via its cation π-interaction, too. One example is the interaction between neurotransmitters and their cognate neuroreceptors with several cation π-interactions involved (Dougherty, 2007). In Crouch *et al.* a critical involvement of a phenylalanine from surfactant protein D in binding with glycoconjugates was found (Crouch *et al.*, 2006). This indicated that, without any charge or functional group located at its side chain, F3 from Gly m 4 might be relevant for binding of IgE antibodies in the context of a Gly m 4-type molecular surrounding provided by the NCS model protein. Same holds true for F5 which was identified as potential functional epitope in combination with D100. Both are close to each other on Gly m 4 surface with F5 located in first and D100 in sixth β-sheet. Aspartate residues with their acidic carboxyl group at their side chain often contribute to protein-protein bindings via ionic interactions. Therefore both residues, F5 and D100, might be important for IgE binding and their combination might even increase IgE signal intensity. But involvement of each single residue in IgE interactions can only be evaluated with NCS variants showing only one grafted putative functional epitope. Together, F3, F5 and D100 are parts of mimotopes identified by Mittag *et al.*, while D100 is also in the region of corresponding Bet v 1 peptides identified by IgE

inhibition experiments performed by Gieras *et al.* (Mittag *et al.*, 2006; Gieras *et al.*, 2011).

Potential functional epitopes for IgE identified in ≥75% of all patients are residues K20, D25/D27, N28/K32, D60 and N128. Concerning these residues, only D60 was not identified with mimotope analysis performed by Mittag *et al.* (Mittag *et al.*, 2006). In addition K20 was already identified as part of a mimotope binding anti-Bet v 1 IgE, while N28 and K32 were mentioned as putative IgE-binding by different analysis with Bet v 1 (Ganglberger *et al.*, 2001; Spangfort *et al.*, 2003; Holm *et al.*, 2004). Several studies using phage-display or monoclonal Bet v 1-directed antibodies identified G61, which is structurally homologous in Bet v 1 to D60 in Gly m 4 as IgE binding (Jensen-Jarolim *et al.*, 1998; Ganglberger *et al.*, 2001; Levin *et al.*, 2014). Recently, using a software tool, N128 was predicted as IgE-binding in Bet v 1-homologous cherry allergen Pru av 1 (Dall'Antonia *et al.*, 2011). Only amino acids D25 and D27 were not described as putative functional epitopes or parts of any known IgE binding patch in literature. Besides several residues already known for their potential IgE binding throughout literature, additional amino acids could be identified as putative functional IgE epitopes. As expected, potential functional IgE epitopes are distributed over complete Gly m 4 surface and can be used as initial results for subsequent analyses (see 4.3.3 and 4.4).

4.3.2 Known BV16 epitope as minor IgE-relevant epitope in Gly m 4

$\Delta 51NCS_{N42/E44/N46/P49/K54}$ was already described as capable of binding BV16 antibody as parts of the known BV16 epitope were grafted onto the molecular surface of this variant. Therefore $\Delta 51NCS_{N42/E44/N46/P49/K54}$ was analyzed for IgE binding together with sera of patients used in this study (see 3.4.1). Not only IgE binding with non-purified cell supernatant but also purified protein was evaluated by immunoblots. Both results are in great accordance whereas only patient 43 showed IgE binding with purified $\Delta 51NCS_{N42/E44/N46/P49/K54}$ but not when using non-purified cell supernatant. For the screening of IgE binding to $\Delta 51NCS_{N42/E44/N46/P49/K54}$ immunoblot was preferred over ELISA analysis because of reduced amounts of allergen and patients' sera needed considering the more descriptive readout derived from ELISA experiments. Since IgE binding to NCS variants only occurs if proteins show Bet v 1-type protein conformation only variants which were able to properly refold upon

denaturing SDS-PAGE and transfer to nitrocellulose membrane are detected. If no detectable signal could be observed even after an increased exposure time of 10 min no binding between $\Delta 51NCS_{N42/E44/N46/P49/K54}$ and serum IgE was concluded. $\Delta 51NCS_{N42/E44/N46/P49/K54}$ showed three grafted residues conserved in both Bet v 1a and Gly m 4 compared to the NCS model protein. E44, N46 and K54 are in close proximity and seem to present an IgE-binding patch, because 41% of patients with and 29% without soy allergy exhibited IgE binding to $\Delta 51NCS_{N42/E44/N46/P49/K54}$. In contrast to a major relevance of these functional epitopes determined in patients with birch-pollen allergy, where 71% showed significant IgE binding, BV16 epitope might be of minor relevance concerning birch-related soy allergy (Berkner *et al.*, 2014). Relevance of BV16 epitope for IgE binding via Bet v 1a was shown by several other publications, too (Mirza *et al.*, 2000; Spangfort *et al.*, 2003; Holm *et al.*, 2004). IgE binding with $\Delta 51NCS_{N42/E44/N46/P49/K54}$ was increased in patients with clinically confirmed soy allergy, indicating higher levels of IgE antibodies directed against BV16 epitope in this group of patients. But in total no correlation between grafted residues and birch-related soy allergy could be drawn, as subjects without soy allergy showed IgE binding to this epitope, too.

4.3.3 Definition of six different IgE epitopes

Regarding published crystal structures of antigen-antibody complexes the area of one allergen epitope was described with 600-900 $Å^2$ (Davies *et al.*, 1990; Braden *et al.*, 1995). Furthermore two independent or distinct epitopes are necessary for simultaneous binding of two antibodies resulting in membrane receptor cross-linking, which represents a crucial step in the pathomechanism of IgE-mediated allergy. Distinct means that two IgE molecules can bind simultaneously to the allergen. Regarding these aspects the identified functional IgE epitopes were clustered into six different putative epitope patches manually (see 3.5). The six patches are located on each side of the protein just like the six faces of a cubical die with an average of 651.2 $Å^2$ (311.4-967.3 $Å^2$) which is in the expected range of 600-900 $Å^2$ for antibody-antigen binding sites. Smallest epitope patch V is below 600 $Å^2$ but presents parts of known BV16 epitope identified by a monoclonal antibody (Berkner *et al.*, 2014). The frequency of six epitope patches based on most frequent amino acid ranged between 56% and 100%. With this arrangement of putative IgE epitope patches it is possible

that different IgE antibodies bind at the same time to Gly m 4, which is a prerequisite for eliciting allergic symptoms. This model also emphasizes that not the presence of one IgE antibody in patient's serum can correlate with clinical soybean allergy, but the combination of different IgE molecules binding to distinct IgE epitopes simultaneously is relevant for an allergic response (cf. 1.1). With additional experiments e. g. RBL release assays this questions might be addressed in more detail. In this context of interaction between different IgEs and a specific allergen the sterical hindrance between large IgE molecules (190 kDa) interacting with small allergens (16.6 kDa) needs to be considered. Binding of an IgE antibody to one epitope on Gly m 4 might prevent binding of a second antibody directed to another epitope in close proximity. Therefore no allergic response is observable concerning these two antibodies although both epitopes are present on Gly m 4 and the patient has increased levels of both IgE antibodies. Sterical hindrance not only occurs between different IgE molecules but also via binding of other immunoglobulins. In this context especially the blocking ability of IgG_4 in allergen-specific immunotherapy is discussed to date (James *et al.*, 2011; Jutel *et al.*, 2013). Therefore a future thorough allergic diagnosis might not only consider allergen-specific IgE levels, but also the allergen-specific epitope repertoire for IgE, as well as levels and epitopes for specific IgG_4, IgM or IgA molecules.

4.4 NSC variants and IgE epitope profile

To further evaluate the identified potential functional IgE epitopes several NCS variants combining multiple functional epitopes were characterized (see 3.5.1). Out of 14 generated variants eleven exhibited native-type secondary structure as revealed by CD spectroscopy. Therefore amino acid substitutions likely did not alter Bet v 1-type protein fold in these variants which is a prerequisite for correlation of grafted residues with IgE binding. $\Delta NCS_{N42/P49}$ variants presented either one individual epitope patch ($\Delta NCS_{N42/P49}$_1-5) or a combination of grafted residues belonging to 2-5 different IgE epitope patches ($\Delta NCS_{N42/P49}$_6-11). With these variants a suitable toolbox for a detailed analysis of Gly m 4-specific IgE epitope profile was generated. To determine the IgE binding of each $\Delta NCS_{N42/P49}$ variant, all patients of groups 1 and 2 were analyzed by dot blotting (see 3.5.2). Dot blot analysis was comparable to IgE-binding assay performed for evaluation of candidate residues with several

differences: i) all $\Delta NCS_{N42/P49}$ variants were purified to homogeneity out of *E. coli* inclusion body fractions rather than used from non-purified cell supernatants, ii) additional step of protein separation via SDS-PAGE was left out and iii) IgE signals could be quantified using densitometry. As expected, all $\Delta NCS_{N42/P49}$ variants were expressed solely in *E. coli* inclusion bodies as already described for $\Delta 51NCS_{N42/P49}$ model protein and $\Delta 51NCS_{N42/P49}$ variants with only single grafted residues. Due to usage of purified proteins showing only one potential specific IgE signal in immunoblot, time-consuming SDS-PAGE could be avoided and $\Delta NCS_{N42/P49}$ variants were blotted directly onto nitrocellulose membrane. As IgE binding requires native protein conformation only NCS variants which were able to properly refold into Bet v 1-type conformation upon transfer to nitrocellulose membrane were detected. Therefore dot blot analysis resembles an easy and fast approach for analysis of eleven $\Delta NCS_{N42/P49}$ variants with all patients' sera compared to ELISA or ImmunoCAP™ for IgE detection where, in addition, markedly increased amounts of patients' sera is needed. Nevertheless IgE ELISA might provide more reliable results or data read-outs compared to dot blot analysis and might be used as additional tool for further IgE-binding analyses using $\Delta NCS_{N42/P49}$ variants. Furthermore each IgE signal was quantified via densitometric analysis. This allows identification of IgE-binding variants in relation to the background signal obtained with $\Delta 51NCS_{N42/P49}$ model protein and allowed exclusion of signals close to background noise.

In total, IgE binding frequencies differed markedly between subjects with and without soybean allergy for several $\Delta NCS_{N42/P49}$ variants. For patients with soy allergy $\Delta NCS_{N42/P49}$_4, 7, 9 and 10 showed highest binding frequencies. In the group of patients without soybean allergy $\Delta NCS_{N42/P49}$_4, 7 and 10 were identified most frequently. Comparable frequencies (≤20% difference) were observed in both groups concerning variants $\Delta NCS_{N42/P49}$_1, 2, 3, 4, 5, 7, 10 and 11. $\Delta NCS_{N42/P49}$_7, 10 and 11 present epitopes of combinational IgE patches while in $\Delta NCS_{N42/P49}$_1-5 solely one epitope patch was grafted. This emphasizes again that patients of both group did not differ mainly in their IgE levels at least directed against epitope patches 1-5. A difference in binding frequency of 29% was observed for $\Delta NCS_{N42/P49}$_6. Furthermore $\Delta NCS_{N42/P49}$_1, 3, 5, 8 and 11 showed different IgE-binding pattern as IgE binding was observable in 1-7 of patients with and in none of patients without soybean allergy. A large difference was observed with $\Delta NCS_{N42/P49}$_9 showing an IgE binding frequency of 68% (15/22) in patients with and only 7% (1/14) in patients without

birch-related soy allergy. Therefore residues of this epitope might be more relevant in the group of subjects with clinically confirmed soybean allergy. In dot blot analysis, $\Delta NCS_{N42/P49}$_9 yields a sensitivity of 68% and a specificity of 93% resulting in positive (PPV) and negative predicted values (NPV) of 94% and 65%, respectively, at least with the study population and the NCS control protein used in this study. High values for both specificity and PPV indicate a potential use of $\Delta NCS_{N42/P49}$_9 as biomarker for soy allergy. With $\Delta NCS_{N42/P49}$_9 a differentiation between patients with soybean allergy and patients only sensitized to soy might be possible. Reference protein $\Delta 51NCS_{N42/P49}$ showed weak IgE binding with sera of several patients in both, initial immunoblot screening and dot blot analysis, most likely because of identical amino acids present in $\Delta 51NCS_{N42/P49}$ and Gly m 4 or Bet v 1a primary sequence with potential importance for interaction with IgE antibodies. In total $\Delta 51NCS_{N42/P49}$ shares 31 and 35 residues with Gly m 4 and Bet v 1a, respectively, whereas 15 and 16 amino acids might be involved in IgE binding because of their side chain characteristics (E, F, H, K, N, S, T, Y).

Again, as already stated in initial screening (see 3.4), patients of both groups can show a low or a high number of identified $\Delta NCS_{N42/P49}$ variants binding patients' IgEs. In total, patients with one single positive IgE signal but also patients with IgE binding to at least 5 out of 11 $\Delta NCS_{N42/P49}$ variants were identified in both groups of patients. Nevertheless, as already shown for screening of NCS library, no group-specific binding pattern could be found in dot blot analysis, too. Rather a patient-specific IgE epitope profile is supposed, as majority of patients exhibited individual binding patterns. Only patients without soybean allergy showed comparable IgE binding patterns. Nevertheless most frequently identified variants $\Delta NCS_{N42/P49}$_4, 7 or 10 in this group of patients do not allow a differentiation between patients with and without soy allergy, as these variants were identified as IgE binding with comparable frequencies in the group of patients with soybean allergy as well. Only $\Delta NCS_{N42/P49}$_9 presents a potential marker for soybean allergy and emphasizes that epitope-resolved *in vitro* diagnosis might allow a more accurate correlation with clinical symptoms compared to component-resolved diagnosis using ImmunoCAP™ for example. As no statistical analyses were conducted within my study population the usability of this system needs to be analyzed statistically in further studies.

Overall dot blot results were in accordance with results of initial screening and emphasize the need of a more patient-specific diagnosis to evaluate presence and

symptomatology of birch-related soybean allergy. Only with $\Delta NCS_{N42/P49}$_9 a correlation between IgE binding and clinically confirmed soybean allergy was observable in majority of patients. Despite of this, only by measuring IgE binding to identified IgE epitopes of Gly m 4, no conclusions on clinical relevance and severity could be drawn. Furthermore number of identified epitope patches or functional IgE epitopes does not necessarily correlate with appearance of soybean allergy or specific symptoms in this study population. Nevertheless, determination of IgE-binding epitopes in each patient might predict relevant combinations of epitopes for birch-related soy allergy. For example, binding of an antibody directed against epitope I prevents interaction of the allergen with IgE via epitope VI due to a potential sterical hindrance (cf. Figure 13). Analyzing these patient-specific selective combinations e. g. by RBL cell release experiments, might give additional information on IgE epitope profile of Gly m 4. Comparable analyses were performed by Christensen *et al.* showing that natural occurring isoallergen variations as well as total and specific IgE repertoire results in different IgE binding affinities (Christensen *et al.*, 2008; Christensen *et al.*, 2010).

In summary, dot blot results with purified $\Delta NCS_{N42/P49}$ variants were more specific for diagnosis of birch-related soy allergy compared to the initial screening using NCS variants but also ImmunoCAP™ analysis. With both, ImmunoCAP™ and initial screening of variants, no correlation between the two groups of patients, their individual symptoms or the severity of symptomatology with identified functional IgE epitopes or specific IgE level was possible. Only by using dot blot with $\Delta NCS_{N42/P49}$ variants presenting characteristic epitope patches a possible correlation between IgE signals and presence of soybean allergy was observed. In this regard $\Delta NCS_{N42/P49}$_9 was identified as potential marker for soybean allergy. Component-resolved diagnosis like ImmunoCAP™ measurements but also initial screening of NCS variants only revealed sensitization profiles of patients to specific recombinant allergens but no correlation between sensitization and clinical relevance in birch-related soybean allergy could be drawn. This is in line with several studies where no predictions on clinical outcome of an individual patient were possible using component-resolved diagnosis (Reuter *et al.*, 2006; Hansen *et al.*, 2009; Ebo *et al.*, 2010; Ballmer-Weber *et al.*, 2012; Guhsl *et al.*, 2015). According to this, the usage of an epitope library might represent a potential tool to differentiate between birch pollen-allergic subjects with and without clinically confirmed birch-associated allergy

to soybean or further related allergens. Nevertheless no correlation between IgE binding pattern and specific symptoms or severity of allergic reactions was found in dot blot analysis.

4.5 Evaluating Gly m 4 specificity of identified IgE epitopes

By analyzing binding of serum IgEs from a pool of patients with and without soybean allergy with NCS variants a specificity to Gly m 4 was found (see 3.5.3). Using eleven generated $\Delta NCS_{N42/P49}$ variants a reduction of Gly m 4-specific IgE binding of at least 70% was observed (cf. Figure 20). This illustrates that identified IgE epitopes and grafted residues are specific for interaction with IgE antibodies present in patients with sensitization to Gly m 4. Self-inhibition with rGly m 4 resulted in a reduction to 2.7 ± 2.1% and 7.6 ± 3.0% of the IgE binding in patients with and without soybean allergy, respectively. Furthermore these results suggest that the vast majority of IgE epitopes of Gly m 4 are covered by our model protein-based Gly m 4-specific epitope library containing a total number of 25 amino acid substitutions towards Gly m 4. Differences in IgE inhibition using $\Delta NCS_{N42/P49}$ variants and rGly m 4 might be caused by an additional IgE epitope present on Gly m 4 and not identified via NCS analysis. This not identified epitope might be the cause for remaining IgE binding of 20.0 ± 5.3% and 30.0 ± 5.3% in pools depleted with NCS variants while only 2.7 ± 2.1% and 7.6 ± 3.0% of IgE binding was observed with rGly m 4-depleted IgE antibodies. Further explanation might be a decreased affinity in binding between IgEs and NCS variants compared to IgEs and Gly m 4 in both groups. Such a decrease might be caused by an incomplete transfer of IgE epitopes from Gly m 4 to NCS. In addition increased concentrations of inhibitors might result in full inhibition especially using rGly m 4 while in this experimental set-up only single point inhibitions were performed. Characteristic residues of each identified IgE epitope were grafted from Gly m 4 to NCS but several amino acids additionally contributing to allergen-antibody interaction are still missing. These residues might also be relevant for binding IgE and therefore cause a decreased antibody affinity. As only 25 residues grafted from Gly m 4 to $\Delta 51NCS_{N42/P49}$ model protein are analyzed in these depletion experiments a increased inhibition might be observed using additional $\Delta 51NCS_{N42/P49}$ variants presenting further identified functional IgE epitopes. $\Delta 51NCS_{N42/P49}$ model protein lead to a minor reduction of Gly m 4-specific binding to IgE as well. Again, as already

discussed (see 4.4), this indicates the presence of several residues in original NCS sequence with IgE-binding capability.

4.6 Functional IgE epitopes identified with rGly m 4 variants

In addition to epitope analysis using NCS library further studies with rGly m 4 and five substitutional rGly m 4 variants were performed (see 3.6). In contrast to NCS system where grafting of functional epitopes lead to an increase in IgE signal, rGly m 4 variants with critical substitutions showed a reduced binding of antibodies. In this study, five variants of rGly m 4 were generated with four to 14 amino acid substitutions belonging to 2-6 of the identified putative IgE epitope patches (see 3.6.2). With this, a total number of 18 residues were addressed. Most residues were substituted to alanine not relevant for the initiation of antibody binding due to the absence of any functional group or charged residues. Four residues, E44, N77, N108 and K133, showed conservative substitutions to serine, lysine or glutamate representing residues with contradictory characteristics than original amino acids, already described for respective residues in Bet v 1 (Spangfort *et al.*, 2003; Holm *et al.*, 2004). rGly m 4_9x was based on rGly m 4_4x variant and rGly m 4_11x and rGly m 4_14x both included substitutions of rGly m 4_10x variant. CD spectra of rGly m 4 variants were comparable to native Gly m 4 and typical for PR10-proteins indicating a Gly m 4-type secondary structure with minor deviations due to the amino acid substitutions. Nevertheless a tendency to unstructured proteins was observable with increasing number of substitutions. Native-type protein folding is essential for studying influence of amino acid substitutions on IgE binding because an altered protein conformation often reduces binding of serum IgE (Holm *et al.*, 2004; Wallner *et al.*, 2011). To study the impact of substituted residues on IgE binding, purified proteins were used together with patients' sera (see 3.6.3). Instead of being expressed in both soluble fraction and *E. coli* inclusion bodies, as shown for rGly m 4 (see 3.6.1), the corresponding variants showed target protein expression for the most part in bacterial inclusion bodies. Therefore introduced amino acid substitutions influenced protein expression and caused an increase in formation of inclusion bodies. rGly m 4 and its variants showed high purity with some minor degradation products according to SDS-PAGE.

Western blot and inhibition ELISA were used to determine IgE binding of rGly m 4

variants with serum IgE (see 3.6.3). In both assays a reduction in IgE signals compared to native Gly m 4 was observable. Majority of patients (94%, 15/16) analyzed with rGly m 4_11x in immunoblot showed a reduced IgE binding, while only one patient with comparable IgE signals could be identified. At least some of substituted residues might represent amino acids crucial for IgE binding and are part of one or more epitopes on Gly m 4. Nevertheless further epitopes need to be present because, as expected, eleven substitutions did not reduce IgE signal completely in all patients. Especially patient 11, not responding to any of introduced residues might contain IgE antibodies directed solely against epitopes different from ones addressed with rGly m 4_11x. As IgE binding requires a native protein conformation all rGly m 4 variants were able to properly refold into their native conformation upon transfer to nitrocellulose membrane and possible differences in refolding behaviors between the rGly m 4 variants were addressed via ELISA experiments where soluble allergens are used. In ELISA rGly m 4_4x showed only minor differences in IgE binding compared to native Gly m 4 indicating that substituted residues T93, E95, K96 and E120 did not present putative functional IgE epitopes or at least their effect on IgE signal is very low compared to overall binding of serum IgEs. This is in line with initial screening of NCS library were $\Delta 51NCS_{N42/P49/E95/K96}$ was one out of two NCS variants not binding IgE in any of used serum. In contrast a decrease in IgE binding was observed with rGly m 4_9x, 10x, 11x and 14x as inhibitors. Often with increasing number of substitutions a stronger decrease in IgE signals was observable. Additional information on IgE inhibition potency of used variants might be achieved using dose-related inhibitions rather than single inhibitor concentrations but substituted residues contributed at least to some extend to binding of IgEs. In none of studied patients a complete reduction of IgE binding was possible with any of the used inhibitors. This emphasizes the presence of further IgE epitopes still allowing interaction with IgE antibodies. Again, as already shown with NCS-based analysis, no correlation between identified IgE epitope pattern and a group of patients or specific symptoms was found. Rather, this approach again suggests a patient-specific IgE epitope profile. Most of 18 substituted residues were identified as putative functional IgE epitopes with NCS analysis, too. Only N77 and Y149, which were not included as candidate residues in initial screening of NCS library had potential IgE binding in rGly m 4 variants. NCS variants $\Delta 51NCS_{D25/D27/N42/P49}$, $\Delta 51NCS_{N28/K32/N42/P49}$ and $\Delta 51NCS_{N42/P49/D60}$ showing grafted

residues D27, K32 and D60 bound serum IgE of at least 75% of all sera in initial screening. K32 and D60 were also substituted in rGly m 4_9x whereas all three residues were changed to alanine in rGly m 4 variants 10x, 11x and 14x. Therefore D27, K32 and D60 might play a key role in binding of IgE antibodies probably due to their surface-exposed position or their charged side chains enabling for easy interaction with antibody ligands.

Overall, both assays indicated that usage of rGly m 4 variants with PR10-type protein conformation is a suitable approach to identify residues contributing to IgE binding. In this study 18 putative functional IgE epitopes could be identified but usage of multiple substitutional rGly m 4 variants requires further experiments to verify importance of each single amino acid. Nevertheless, generated variants can be used anyway as a basis for thorough studies on IgE epitope profile.

4.7 Analysis of IgE antibody avidity in study population

In addition to IgE repertoire and IgE epitope profile, differences in antibody affinity and avidity are discussed as potential reason for clinical reactions to allergens. Different IgE antibodies might bind epitopes on allergen surface with a range of affinities. Several high affinity epitopes are for example known for peanut allergen Ara h 3 and melon allergen Cuc m 2 (Rougé *et al.*, 2009; Tordesillas *et al.*, 2010). In this study, both groups of patients with and without birch-related soybean allergy were analyzed according to differences in IgE avidity (see 3.7 and 3.8.3). For this, binding between IgE and rGly m 4 was inhibited with rGly m 4 and the misfolded variant rGly m $4_{S111P/L150P}$. Overall no differences in avidity were observable between both groups of patients regarding each allergen. Using rGly m 4 a complete inhibition in IgE binding was possible. rGly m $4_{S111P/L150P}$ was able to inhibit binding of IgE to rGly m 4 with >80% in both groups as well. This is in line with analysis of misfolded rGly m $4_{S111P/L150P}$ exhibiting IgE binding with several patients of study population. Due to less patients showing IgE binding with rGly m $4_{S111P/L150P}$ compared to rGly m 4 in immunoblot, a shift in IgE inhibition to higher concentrations of misfolded variant was observable in ELISA. For rGly m 4 an increase in IgE avidity might not be relevant for outbreak of soybean allergy in the study population. Rather affinity of specific IgE epitopes should be considered in combination with patient-specific IgE epitope profile. High affinity epitopes causing allergic symptoms might be present

only in patients with clinical allergy while patients sensitized but not allergic to soy show no immunodominant epitopes. Therefore thorough analysis of IgE epitopes and their antibody affinities might be useful to determine allergic potential.

4.8 Misfolded variants of Gly m 4 and Bet v 1a

Recombinant preparations of Gly m 4 and especially Bet v 1a are used in both diagnostic and scientific applications. In basic research rBet v 1a is often used as reference material for biological and physicochemical analysis with its well established purification protocol also existing for rGly m 4. Furthermore both allergens are used in *in vitro* diagnostics for example in ImmunoCAP™ measurements to determine allergen-specific IgE or IgG amounts in patients' sera. Recombinant Bet v 1a is available as a reference standard preparation from the European Directorate for the Quality of Medicines and Health Care (EDQM). For diagnostic purpose and laboratory use a consistent quality of both allergen preparations needs to be maintained. In this study I analyzed the physicochemical and immunological characteristics of both rBet v 1a and rGly m 4 and the influence of different molar ratios of a respective misfolded variant concerning structure and IgE binding ability (see 3.8). With this, a correlation between experimental and theoretically expected data for both allergen combinations is possible. Furthermore the precision according to the confidence intervals of 95% chosen with the actual IgE binding protein moiety is resolved quantitatively by the respective analytical method. The variants generated of both, Bet v 1a and Gly m 4, showed two substitutions to proline located in seventh β-sheet and C-terminal α-helix, respectively (see 3.8.1). Substitution S112P in Bet v 1a and its homologue Mal d 1 in apple was already characterized by Son *et al.* as a protein variant with decreased IgE binding capacity and drastically reduced protein folding (Son *et al.*, 1999). Hence, these observations were also expected for Gly m 4. An additional substitution to proline was introduced in C-terminal α-helix of both proteins to prevent formation of this secondary structure element. Resulting variants rBet v $1a_{S112P/R145P}$ and rGly m $4_{S111P/L150P}$ showed an unfolded protein structure in CD and NMR spectroscopy (see 3.8.2). Both were analyzed for allergen-antibody interaction via immunoblot and as IgE binding requires native protein conformation only allergens which were able to properly refold into Bet v 1-type conformation upon SDS-PAGE and subsequent transfer to nitrocellulose

membrane were detected (see 3.8.3). Misfolded variant rBet v 1a$_{S112P/R145P}$ was not able to bind IgE antibodies of patients sensitized to Bet v 1a or birch pollen. In contrast, in subjects with sensitization to Bet v 1a and Gly m 4 a binding between IgE and at least one misfolded variant was possible in the majority (92%, 33/36 subjects) of patients analyzed. Thereby no difference between subjects with or without clinical reactivity to soybean or specific symptoms could be observed. Binding characteristics of rBet v 1a$_{S112P/R145P}$ and rGly m 4$_{S111P/L150P}$ did not correlate with clinical soybean allergy or any allergic symptoms. If no detectable signal could be observed even after an increased exposure time of 10 min no binding between misfolded variant and serum IgE was concluded, but usage of a quantitative readout of IgE-binding might result in more detailed conclusions. The remaining potential of interaction between misfolded variants and IgE antibodies suggests the presence of at least one linear IgE epitope in Bet v 1a$_{S112P/R145P}$ and Gly m 4$_{S111P/L150P}$ sequence. This linear stretch of adjacent amino acids might be IgE binding in both native allergens or became accessible due to protein unfolding and shows no IgE binding in both well-folded rBet v 1a and rGly m 4. A further explanation for IgE binding with both misfolded variants might be the presence of at least one intact secondary structural element forming a linear or conformational IgE epitope. The structural element which is potentially maintained is however not detectable in both CD and NMR spectroscopy indicating its small ratio in comparison to total unfolded protein structure. The presence of a very low amount of fully folded material in the misfolded variant can be excluded as no IgE binding could be detected with several sera and the identical batch of purified misfolded variant. To answer this question, further investigations are required. For example using RBL cell release assay it would be possible to analyze if only one or at least two distinct epitopes are present. Only presence of two or more epitopes allows cross-linkage of IgE antibodies binding to different epitopes causing release of mediator molecules. If no release would be detected this emphasizes that only one single IgE epitope is left on Gly m 4. Analysis of ten birch-sensitized patients which did not show IgE binding with rBet v 1a$_{S112P/R145P}$ indicates the absence of antibodies directed against the linear or conformational epitope remaining in misfolded variant at least in this study population.

Analysis of different molar ratios of rBet v 1a/rBet v 1a$_{S112P/R145P}$ and rGly m 4/rGly m 4$_{S111P/L150P}$ in CD spectroscopy revealed comparable results (see 3.8.4). With increasing amount of misfolded variant a shift from Bet v 1-type CD

spectrum to spectra characteristic for unstructured proteins was observable. In immunoblots both allergens showed a successive reduction in IgE binding from native allergen to their respective misfolded variant. Both methods are well-suited to trace protein characteristics of native allergens to misfolded variants showing unstructured secondary structures and no/low IgE binding.

Statistical analysis were only performed for rBet v 1a/rBet v $1a_{S112P/R145P}$ ratios where three independent preparations of both proteins were generated resulting in three pairs with incrementally increasing ratios of rBet v $1a_{S112P/R145P}$. In immunoblot, CD spectroscopy, ELISA and RBL mediator release assay mixtures with <10% of rBet v 1a showed large differences between measured and theoretical values. In this range CD, ELISA and mediator release assay underestimated the ratio of rBet v 1a while in immunoblotting an overestimation was found. In contrast, all assays showed good correlations from 100-90% rBet v 1a. Furthermore the 95% confidence intervals of each assay were determined. CD spectroscopy and ELISA performed comparable where a reduction from 100% to 60% rBet v 1a revealed an accuracy of 1.03 and 0.93, respectively. In contrast, in immunoblot and RBL mediator release assay 20% or even only 1% rBet v 1a could be differentiated in 95% confidence interval from total IgE binding with 100% rBet v 1a. Statistical significance was found in all assays except RBL mediator release assay, probably because of its biological background. A significant reduction in secondary structure or IgE binding was observable at 60% rBet v 1a in CD spectroscopy, 80% and 20% rBet v 1a in both immunological assays which are immunoblot and ELISA. Overall obtained results of both CD spectroscopy and ELISA correlated best with theoretically expected values at least for the range between 100-10% rBet v 1a, making them most suitable for determination of allergen concentration. Larger differences appeared when analyzing actual/theoretical data from immunoblot as well as RBL mediator release assay. These assays required large differences in concentrations of allergen and misfolded variant for identification of IgE binding moiety. Taken together, all assays used in this study are well-suited to trace changes in ratios of active and misfolded variant or detect even weak IgE signals in protein samples. However the list of applied methods is not complete in this study and further applications should be considered to detect traces of misfolded variants in allergen preparations. Using surface plasmon resonance, isothermal titration calorimetry, infrared spectroscopy or usage of the recently identified physiological ligand of Bet v 1a might represent reasonable methods to analyze the

quality of allergen preparations (Seutter von Loetzen *et al.*, 2014). IgE binding of allergen preparations might also be influenced by additional effects like storage conditions or modifications in amino acid side chains. Taken together the misfolded variants rBet v $1a_{S112P/R145P}$ and rGly m $4_{S111P/L150P}$ might be used as quality reference standards for example in screening of hypoallergenic molecules with potential use in treatment of allergies using several applications to measure IgE binding capacity and quality of allergen preparations.

5 SUMMARY

Individuals with birch pollinosis may show allergic reactions after consumption of soybean-containing food. This is caused by cross-reaction of IgE directed against the major birch pollen allergen Bet v 1 with the structurally homologous allergen Gly m 4 from soybean. Hypersensitivity reactions in birch-soy allergy range from mild reactions to severe systemic reactions. Sera of birch pollen-allergic subjects may contain IgE to Gly m 4, even though no allergy to soy is present. Thus Gly m 4-specific IgE *per se* is not a suitable biomarker for birch-related soy allergy. To develop novel approaches for improved diagnosis and therapy of birch-soy allergy, knowledge on epitopes and IgE epitope profile of Gly m 4 for is needed. To date, data on epitopes of Bet v 1 and its homologous allergen Gly m 4 is very limited.

In this study, birch pollen-allergic patients with (27 subjects) and without (20 subjects) clinically confirmed allergy to soybean were included and analyzed regarding Gly m 4-specific serum IgE/IgG levels and epitope profiles of Gly m 4 for IgE. Specific IgE levels against Bet v 1, Gly m 4 and further soy allergens Gly m 5 and Gly m 6 were determined by ImmunoCAP™ and IgE binding to rGly m 4 and soy extract was tested in western blot. To analyze putative IgE epitopes of Gly m 4, non-allergenic Norcoclaurine synthase (NCS) from meadow rue was used as a model protein. NCS is structurally homologous to Gly m 4 but exhibits none to very little binding of Gly m 4-specific IgE antibodies enabling grafting of Gly m 4 epitopes onto the model protein. Potential candidate residues of Gly m 4 were selected by bioinformatic analysis of antibody-binding phage-displayed peptides and mapping of segments of Gly m 4 primary structure onto the molecular surface of the allergen. As a result a library of recombinant NCS variants with potential IgE binding was generated. In addition, five multiple substitutional variants of rGly m 4 with a total number of 18 amino acid substitutions crucial for IgE binding were generated and analyzed for antibody binding with patients' sera. Furthermore a misfolded variant of each rBet v 1a and rGly m 4 was generated and defined molar ratios of folded/misfolded variants were compared in different immunological and physicochemical assays.

Gly m 4-specific median IgE and IgG levels of allergic (IgE: 9.3 kU_A/L, IgG: 8.1 mg_A/L) and non-allergic (IgE: 4.5 kU_A/L, IgG: 8.3 mg_A/L) subjects were comparable. The specific IgE levels did not correlate to the (severity of) clinical phenotypes. 51 candidate residues of Gly m 4 were selected for IgE epitope analysis

and 46 potential functional IgE epitopes, single residues within a structural IgE epitope which dominate the energetics of allergen-IgE binding, were identified with IgE binding to $\Delta NCS_{N42/P49}$ variants. The putative functional IgE epitope pattern was individual for each patient and not distinguishable between allergic and non-allergic subjects. Using five rGly m 4 variants parts of the results of the NCS-based analyses could be confirmed with 18 potential functional IgE epitopes identified. 46 potential functional IgE epitopes clustered into six distinct putative IgE-binding areas on Gly m 4 and eleven NCS variants ($\Delta NCS_{N42/P49}$_1-11) presenting parts of these epitope areas were purified, showing a Gly m 4-type secondary structure according to CD measurements. Densitometric analysis for binding of IgE antibodies was performed via dot blot with sera of the study population but no characteristic IgE epitope pattern could be found. In contrast $\Delta NCS_{N42/P49}$_9 was identified as most suitable marker to distinguish soy allergic from tolerant patients in birch-related soybean allergy with a sensitivity and specificity of 68% and 93%, respectively. Using a pool of $\Delta NCS_{N42/P49}$ variants a depletion of Gly m 4-specific IgE binding to about 80% and 70% in pooled sera of patients with and without soybean allergy, respectively, was possible. Therefore identified putative functional IgE epitopes are specific for interaction with IgE antibodies in study population. With $\Delta NCS_{N42/P49}$_9 a differentiation between sensitization to Gly m 4 and clinically confirmed soybean allergy might be possible. The usage of a Gly m 4-specific epitope library might be a more promising tool for evaluating birch-related soybean allergy compared to ImmunoCAP™ and screening of substitutional Gly m 4 variants alone. However no correlations between patient's IgE epitope profile and specific symptoms to soy or severity of allergic reactions could be found using NCS-based epitope library. Rather patient-specific epitope pattern might be relevant for birch-related soybean allergy. Therefore a thorough diagnosis by the combination of component-resolved diagnosis and profound epitope analysis might be mandatory in the future.

Using two misfolded variants of rBet v 1a and rGly m 4 the impact of unstructured allergens in rBet v 1a/rGly m 4 preparations was addressed with physico- and immunological assays. Both rBet v $1a_{S112P/R145P}$ and rGly m $4_{S111P/L150P}$ showed a highly disordered protein conformation and reduced IgE binding frequencies with analyzed patients' sera. With CD spectroscopy, immunoblot, ELISA and RBL cell release assay defined combinations of native and unstructured allergen were assessed concerning secondary structure and IgE binding. Correlation of rBet v 1a

content with secondary structure and IgE binding was suitable only at high rBet v 1a$_{S112P/R145P}$ levels in mixtures. CD spectroscopy and ELISA performed more precise compared to immunoblot and rat basophil cell mediator release assay where larger deviations between native and unstructured allergen were necessary. In addition, quantification of IgE-binding allergen was difficult for concentrations of rBet v 1a ≤10% in all assays. Overall, CD, ELISA and RBL cell release assay underestimated while immunoblot overestimated the actual level of rBet v 1a. Results of both misfolded variants rBet v 1a$_{S112P/R145P}$ and rGly m 4$_{S111P/L150P}$ might be used within screening of hypoallergenic molecules with potential use in treatment of allergies or in quality assessment of recombinant allergen preparations.

6 DEUTSCHE ZUSAMMENFASSUNG

Patienten mit einer Birkenpollenallergie zeigen oftmals auch allergische Reaktionen auf sojahaltige Nahrungsmittel oder Getränke. Grund ist eine Kreuzreaktion von IgE-Antikörpern gerichtet gegen das Hauptallergen der Birke, Bet v 1, und dem strukturhomologen Allergen Gly m 4 der Sojabohne. Die klinischen Symptome können in der Birkenpollen-assoziierten Sojaallergie bis hin zu schweren Reaktionen, wie einem anaphylaktischen Schock, reichen. Patienten mit einer Birkenpollen-assoziierten Sojaallergie weisen demnach sowohl IgE-Antikörper gegen Bet v 1 als auch Gly m 4 auf. Auffällig ist hierbei, dass Birkenpollenallergiker ohne assoziierte Sojaallergie oft auch diese beiden Allergen-spezifischen IgEs besitzen. Deshalb kann die Präsenz von Gly m 4-gerichteten IgEs an sich nicht als geeigneter Marker für eine Sojaallergie eingesetzt werden. Um eine exakte Diagnose und wirkungsvolle Therapie zur Behandlung solcher Allergien zu ermöglichen, sind neben Patientendaten auch detaillierte Informationen über die IgE-bindenden Bereiche, die sogenannte Epitope, auf Gly m 4 nötig. Bis heute ist jedoch nur wenig über IgE-Epitope von Bet v 1 und seinem Homolog Gly m 4 bekannt.

In diese Arbeit wurden 27 bzw. 20 Birkenpollenallergiker mit bzw. ohne eine klinisch bestätigte Sojaallergie eingeschlossen und hinsichtlich ihrer Gly m 4-spezifischen IgE- und IgG-Konzentrationen im Serum, sowie IgE-Epitopprofile untersucht. Bet v 1-, Gly m 4-, Gly m 5- und Gly m 6-gerichtete IgE-Level wurden mittels ImmunoCAP™ gemessen und die IgE-Bindung an rGly m 4 sowie Sojaextrakt anhand Western Blots bestimmt. Um mögliche IgE-Epitope analysieren zu können, wurde die nicht-allergene Norcoclaurinsynthase (NCS) aus der gelben Wiesenraute als Modelprotein eingesetzt. NCS ist strukturhomolog zu Gly m 4 zeigt aber nur wenig bis gar keine Bindung an Gly m 4-spezifische IgEs und kann daher für das Epitopgrafting von Gly m 4 auf NCS verwendet werden. Aminosäuren von Gly m 4, die als mögliche IgE-bindende Reste in Frage kommen, wurden mittels einer bioinformatischen Analyse von IgE-bindenden Peptiden einer Phagenbibliothek sowie durch das Mappen von Primärstrukturabschnitten von Gly m 4 auf dessen Proteinoberfläche identifiziert. So konnte eine Bibliothek von rekombinant hergestellten NCS Varianten mit potenzieller IgE-Bindung erzeugt werden. Außerdem wurden fünf Varianten von rGly m 4, die insgesamt 18 potentiell IgE-relevante Aminosäureaustausche zeigen, erzeugt und hinsichtlich ihrer Bindung an

Patienten-IgE untersucht. Zusätzlich wurden zwei ungefaltete Varianten von rBet v 1a und rGly m 4 hergestellt und in unterschiedlichen Mischungen mit korrekt gefaltetem rBet v 1a/rGly m 4 anhand diverser immunologischer und physiko-chemischer Verfahren untersucht.

Hinsichtlich ihrer mittleren Gly m 4-spezifischen IgE- und IgG-Konzentrationen konnte kein Unterschied zwischen Patienten mit (IgE: 9.3 kU_A/L, IgG: 8.1 mg_A/L) und ohne (IgE: 4.5 kU_A/L, IgG: 8.3 mg_A/L) Sojaallergie gefunden werden. Weiterhin korrelierten diese spezifischen IgE-Konzentrationen nicht mit dem klinischen Phänotyp der Patienten. Insgesamt konnten aus 51 Gly m 4-spezifischen Aminosäuren 46 potentiell funktionale IgE-Epitope identifiziert werden. Bei funktionalen Epitopen handelt es sich um einzelne Aminosäuren, welche innerhalb eines strukturellen IgE-Epitops die Energetik der Allergen-IgE-Bindung dominieren. Das potentielle, funktionale IgE-Epitopprofil jedes Patienten war individuell und nicht charakteristisch für die Unterscheidung zwischen Allergikern und Nicht-Allergikern. Durch die Analyse mittels der fünf Substitutionsvarianten von rGly m 4 konnten Teile der Ergebnisse der NCS-Epitopanalyse bestätigt werden. Insgesamt ergaben sich hier 18 potentielle, funktionale IgE-Epitope. Alle 46 potentiell funktionalen IgE-Epitope konnten in sechs mögliche Epitopbereiche auf Gly m 4 zusammengefasst werden. Elf NCS-Varianten ($\Delta NCS_{N42/P49}$_1-11), die Teile dieser Epitopbereiche aufweisen, wurden hergestellt und zeigten zu Gly m 4 vergleichbare CD-Spektren. Eine densitometrische Analyse der IgE-Bindung in Patientenseren wurde mittels eines Dot Blots durchgeführt, jedoch konnte kein charakteristisches IgE-Epitopprofil ermittelt werden. Hingegen konnte die NCS Variante $\Delta NCS_{N42/P49}$_9 mit einer Sensitivität bzw. Spezifität von 68% bzw. 93% als vielversprechender Marker identifiziert werden, um Sojaallergiker von toleranten Patienten zu unterscheiden. Weiterhin war mit einer Mischung aus allen $\Delta NCS_{N42/P49}$ Varianten eine Depletion der Gly m 4-spezifischen IgE-Bindung in den Seren der Patienten mit bzw. ohne Sojaallergie auf etwa 80% bzw. 70% möglich. Demnach sind die identifizierten potentiell funktionalen IgE-Epitope spezifisch für die Bindung von IgE-Antikörpern in der Studiengruppe. $\Delta NCS_{N42/P49}$_9 scheint die passendste Variante zu sein, mit der eine Unterscheidung zwischen Sensibilisierung auf Gly m 4 und klinisch bestätigter Sojaallergie möglich ist. Der Einsatz einer Gly m 4-spezifischen Epitopbibliothek kann somit eine vielversprechendere Möglichkeit zur Analyse von Birkenpollen-assoziierter Sojaallergie, verglichen zu ImmunoCAP™ oder der Untersuchung von Gly m 4-

Substitutionsvarianten darstellen. Jedoch konnte auch mit der NCS-Epitopbibliothek keine Korrelation zwischen dem IgE-Epitopprofil der Patienten und ihren allergischen Symptomen sowie dem Schweregrad der Symptome ermittelt werden. Vielmehr scheint ein patientenspezifisches Epitopprofil Grundlage für eine Birkenpollen-assoziierte Sojaallergie zu sein. Somit könnte also eine Kombination aus rekombinant hergestellten Allergenen und der detaillierten Untersuchung mittels einer Epitopbibliothek zukünftig zu einer genaueren Diagnose führen.

Mittels der beiden Allergenvarianten von rBet v 1a und rGly m 4 konnte der Einfluss von ungefaltetem Protein durch physikochemische und immunologische Methoden in Präparaten aus rBet v 1a bzw. rGly m 4 untersucht werden. Beide Varianten rBet v $1a_{S112P/R145P}$ und rGly m $4_{S111P/L150P}$ zeigten eine ungefaltete Proteinstruktur sowie eine reduzierte IgE-Bindung in den eingesetzten Patientenseren. In CD-Spektroskopie, Immunoblot, ELISA und RBL-Zellassay wurden definierte Mischungen aus nativem und ungefaltetem Allergen hinsichtlich Sekundärstruktur und IgE-Bindung untersucht. Eine Korrelation von rBet v 1a-Konzentration mit der Sekundärstruktur des Allergens sowie der IgE-Bindung war nur bei hohen Konzentrationen an rBet v $1a_{S112P/R145P}$ möglich. Hier lieferten CD-Spektroskopie und ELISA präzisere Ergebnisse als der Immunoblot und RBL-Zellassay. Bei letzteren waren größere Unterschiede zwischen nativem und ungefaltetem Allergen für eine eindeutige Korrelation nötig. Weiterhin war die Bestimmung an IgE-bindendem Allergen in allen Assays für rBet v 1a-Konzentrationen ≤10% schwierig. CD-Spektroskopie, ELISA und RBL-Zellassay führten zu einer Unterschätzung des tatsächlichen Gehalts an rBet v 1a, wohingegen mit dem Immunoblot mehr rBet v 1a gemessen wurde als tatsächlich in der Probe vorhanden war. Die Ergebnisse der beiden Faltungsvarianten rBet v $1a_{S112P/R145P}$ und rGly m $4_{S111P/L150P}$ könnten für das Screening von hypoallergenen Molekülen zur möglichen Behandlung von Allergien oder in der Qualitätskontrolle von Allergenpräparaten eingesetzt werden.

7 REFERENCES

Aalberse, RC, Akkerdaas, J, van Ree, R. 2001. Cross-reactivity of IgE antibodies to allergens. *Allergy* **56**: 478–490.

Aalberse, RC, van der Gaag, R, van Leeuwen, J. 1983. Serologic aspects of IgG4 antibodies. I. Prolonged immunization results in an IgG4-restricted response. *Journal of immunology (Baltimore, Md. : 1950)* **130**: 722–726.

Akdis, CA, Blesken, T, Akdis, M, Wüthrich, B, Blaser, K. 1998. Role of interleukin 10 in specific immunotherapy. *The Journal of clinical investigation* **102**: 98–106.

Akdis, M. 2014. New treatments for allergen immunotherapy. *The World Allergy Organization journal* **7**: 23. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4174392/pdf/1939-4551-7-23.pdf.

Albrecht, M, Kühne, Y, Ballmer-Weber, BK, Becker, W, Holzhauser, T, Lauer, I, Reuter, A, Randow, S, Falk, S, Wangorsch, A, Lidholm, J, Reese, G, Vieths, S. 2009. Relevance of IgE binding to short peptides for the allergenic activity of food allergens. *The Journal of allergy and clinical immunology* **124**: 328-36, 336.e1-6.

Allen, KJ, Koplin, JJ. 2012. The epidemiology of IgE-mediated food allergy and anaphylaxis. *Immunology and allergy clinics of North America* **32**: 35–50.

Amit, AG, Mariuzza, RA, Phillips, SE, Poljak, RJ. 1986. Three-dimensional structure of an antigen-antibody complex at 2.8 A resolution. *Science (New York, N.Y.)* **233**: 747–753.

Asero, R, Ballmer-Weber, BK, Beyer, K, Conti, A, Dubakiene, R, Fernandez-Rivas, M, Hoffmann-Sommergruber, K, Lidholm, J, Mustakov, T, Oude Elberink, Joanne N G, Pumphrey, Richard S H, Stahl Skov, P, van Ree, R, Vlieg-Boerstra, BJ, Hiller, R, Hourihane, JO, Kowalski, M, Papadopoulos, NG, Wal, J, Mills, E N Clare, Vieths, S. 2007. IgE-mediated food allergy diagnosis: Current status and new perspectives. *Molecular nutrition & food research* **51**: 135–147.

Ballmer-Weber, BK, Holzhauser, T, Scibilia, J, Mittag, D, Zisa, G, Ortolani, C, Oesterballe, M, Poulsen, LK, Vieths, S, Bindslev-Jensen, C. 2007. Clinical characteristics of soybean allergy in Europe: a double-blind, placebo-controlled food challenge study. *The Journal of allergy and clinical immunology* **119**: 1489–1496.

Ballmer-Weber, BK, Skamstrup Hansen, K, Sastre, J, Andersson, K, Bätscher, I, Ostling, J, Dahl, L, Hanschmann, K, Holzhauser, T, Poulsen, LK, Lidholm, J, Vieths, S. 2012. Component-resolved in vitro diagnosis of carrot allergy in three different regions of Europe. *Allergy* **67**: 758–766.

Ballmer-Weber, BK, Vieths, S, Lüttkopf, D, Heuschmann, P, Wüthrich, B. 2000. Celery allergy confirmed by double-blind, placebo-controlled food challenge: a clinical study in 32 subjects with a history of adverse reactions to celery root. *The Journal of allergy and clinical immunology* **106**: 373–378.

Ballmer-Weber, BK, Wüthrich, B, Wangorsch, A, Fötisch, K, Altmann, F, Vieths, S. 2001. Carrot allergy: double-blinded, placebo-controlled food challenge and

identification of allergens. *The Journal of allergy and clinical immunology* **108**: 301–307.

Beier, KC, Kallinich, T, Hamelmann, E. 2007. Master switches of T-cell activation and differentiation. *The European respiratory journal* **29**: 804–812.

Berkner, H, Engelhorn, J, Liscombe, DK, Schweimer, K, Wöhrl, BM, Facchini, PJ, Rösch, P, Matecko, I. 2007. High-yield expression and purification of isotopically labeled norcoclaurine synthase, a Bet v 1-homologous enzyme, from Thalictrum flavum for NMR studies. *Protein expression and purification* **56**: 197–204.

Berkner, H, Schweimer, K, Matecko, I, Rösch, P. 2008. Conformation, catalytic site, and enzymatic mechanism of the PR10 allergen-related enzyme norcoclaurine synthase. *The Biochemical journal* **413**: 281–290.

Berkner, H, Seutter von Loetzen, C, Hartl, M, Randow, S, Gubesch, M, Vogel, L, Husslik, F, Reuter, A, Lidholm, J, Ballmer-Weber, BK, Vieths, S, Rösch, P, Schiller, D. 2014. Enlarging the toolbox for allergen epitope definition with an allergen-type model protein. *PLoS ONE* **9**: e111691.

Berman, HM. 2000. The Protein Data Bank. *Nucleic Acids Research* **28**: 235–242.

Bøgh, KL, Nielsen, H, Eiwegger, T, Madsen, CB, Mills, E N C, Rigby, NM, Szépfalusi, Z, Roggen, EL. 2014. IgE versus IgG4 epitopes of the peanut allergen Ara h 1 in patients with severe allergy. *Molecular Immunology* **58**: 169–176.

Bohle, B, Zwölfer, B, Heratizadeh, A, Jahn-Schmid, B, Antonia, YD, Alter, M, Keller, W, Zuidmeer, L, van Ree, R, Werfel, T, Ebner, C. 2006. Cooking birch pollen-related food: divergent consequences for IgE- and T cell-mediated reactivity in vitro and in vivo. *The Journal of allergy and clinical immunology* **118**: 242–249. 16815162.

Bolhaar, STHP, Tiemessen, MM, Zuidmeer, L, van Leeuwen, A, Hoffmann-Sommergruber, K, Bruijnzeel-Koomen, C A F M, Taams, LS, Knol, EF, van Hoffen, E, van Ree, R, Knulst, AC. 2004. Efficacy of birch-pollen immunotherapy on cross-reactive food allergy confirmed by skin tests and double-blind food challenges. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* **34**: 761–769.

Braden, BC, Cauerhff, AN, Dall'Acqua, W, Fields, BA, Goldbaum, FA, Malchiodi, EL, Mariuzza, RA, Poljak, RJ, Schwarz, FP, Ysern, X, Bhat, TN. 1995. Structure and Thermodynamics of Antigen Recognition by Antibodiesa. *Annals of the New York Academy of Sciences* **764**: 315–327.

Breiteneder, H, Ebner, C. 2000. Molecular and biochemical classification of plant-derived food allergens. *Journal of Allergy and Clinical Immunology* **106**: 27–36.

Breiteneder, H, Mills, C. 2006. Structural bioinformatic approaches to understand cross-reactivity. *Molecular nutrition & food research* **50**: 628–632.

Broide, DH. 2001. Molecular and cellular mechanisms of allergic disease. *The Journal of allergy and clinical immunology* **108**: S65-71.

Bublil, EM, Freund, NT, Mayrose, I, Penn, O, Roitburd-Berman, A, Rubinstein, ND, Pupko, T, Gershoni, JM. 2007. Stepwise prediction of conformational discontinuous B-cell epitopes using the Mapitope algorithm. *Proteins* **68**: 294–304.

Bucher, X, Pichler, WJ, Dahinden, CA, Helbling, A. 2004. Effect of tree pollen specific, subcutaneous immunotherapy on the oral allergy syndrome to apple and hazelnut. *Allergy* **59**: 1272–1276.

Bufe, A, Spangfort, MD, Kahlert, H, Schlaak, M, Becker, WM. 1996. The major birch pollen allergen, Bet v 1, shows ribonuclease activity. *Planta* **199**: 413–415.

Burks, AW, Jones, SM, Wood, RA, Fleischer, DM, Sicherer, SH, Lindblad, RW, Stablein, D, Henning, AK, Vickery, BP, Liu, AH, Scurlock, AM, Shreffler, WG, Plaut, M, Sampson, HA. 2012. Oral immunotherapy for treatment of egg allergy in children. *The New England journal of medicine* **367**: 233–243.

Burney, Potts, J, Kummeling, I, Mills, E N C, Clausen, M, Dubakiene, R, Barreales, L, Fernandez-Perez, C, Fernandez-Rivas, M, Le, T, Knulst, AC, Kowalski, ML, Lidholm, J, Ballmer-Weber, BK, Braun-Fahlander, C, Mustakov, T, Kralimarkova, T, Popov, T, Sakellariou, A, Papadopoulos, NG, Versteeg, SA, Zuidmeer, L, Akkerdaas, JH, Hoffmann-Sommergruber, K, van Ree, R. 2014. The prevalence and distribution of food sensitization in European adults. *Allergy* **69**: 365–371.

Bush, RK, Hefle, SL. 1996. Food allergens. *Critical reviews in food science and nutrition* **36 Suppl**: S119-63.

Campbell, DE, Mehr, S. 2015. Fifty years of allergy: 1965-2015. *Journal of paediatrics and child health* **51**: 91–93.

Cavallo, L. 2003. POPS: a fast algorithm for solvent accessible surface areas at atomic and residue level. *Nucleic Acids Research* **31**: 3364–3366.

Chafen, JJS, Newberry, SJ, Riedl, MA, Bravata, DM, Maglione, M, Suttorp, MJ, Sundaram, V, Paige, NM, Towfigh, A, Hulley, BJ, Shekelle, PG. 2010. Diagnosing and managing common food allergies: a systematic review. *JAMA : the journal of the American Medical Association* **303**: 1848–1856.

Chen, WH, Sun, PP, Lu, Y, Guo, WW, Huang, YX, Ma, ZQ. 2011. MimoPro: a more efficient Web-based tool for epitope prediction using phage display libraries. *BMC bioinformatics* **12**: 199.

Christensen, LH, Holm, J, Lund, G, Riise, E, Lund, K. 2008. Several distinct properties of the IgE repertoire determine effector cell degranulation in response to allergen challenge. *The Journal of allergy and clinical immunology* **122**: 298–304.

Christensen, LH, Riise, E, Bang, L, Zhang, C, Lund, K. 2010. Isoallergen variations contribute to the overall complexity of effector cell degranulation: effect mediated through differentiated IgE affinity. *Journal of immunology (Baltimore, Md. : 1950)* **184**: 4966–4972.

Chruszcz, M, Pomés, A, Glesner, J, Vailes, LD, Osinski, T, Porebski, PJ, Majorek, KA, Heymann, PW, Platts-Mills, Thomas A E, Minor, W, Chapman, MD. 2012. Molecular determinants for antibody binding on group 1 house dust mite allergens. *The Journal of biological chemistry* **287**: 7388–7398.

Codina, R, Lockey, RF, Fernández-Caldas, E, Rama, R. 1997. Purification and characterization of a soybean hull allergen responsible for the Barcelona asthma outbreaks. II. Purification and sequencing of the Gly m 2 allergen. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* **27**: 424–430.

Cromwell, O, Häfner, D, Nandy, A. 2011. Recombinant allergens for specific immunotherapy. *The Journal of allergy and clinical immunology* **127**: 865–872.

Crouch, E, McDonald, B, Smith, K, Cafarella, T, Seaton, B, Head, J. 2006. Contributions of phenylalanine 335 to ligand recognition by human surfactant protein D: ring interactions with SP-D ligands. *The Journal of biological chemistry* **281**: 18008–18014.

Crowell, DN, John, ME, Russell, D, Amasino, RM. 1992. Characterization of a stress-induced, developmentally regulated gene family from soybean. *Plant molecular biology* **18**: 459–466.

Dall'Acqua, W, Goldman, ER, Lin, W, Teng, C, Tsuchiya, D, Li, H, Ysern, X, Braden, BC, Li, Y, Smith-Gill, SJ, Mariuzza, RA. 1998. A mutational analysis of binding interactions in an antigen-antibody protein-protein complex. *Biochemistry* **37**: 7981–7991.

Dall'Antonia, F, Gieras, A, Devanaboyina, SC, Valenta, R, Keller, W. 2011. Prediction of IgE-binding epitopes by means of allergen surface comparison and correlation to cross-reactivity. *The Journal of allergy and clinical immunology* **128**: 872-879.e8.

Dall'Antonia, F, Pavkov-Keller, T, Zangger, K, Keller, W. 2014. Structure of allergens and structure based epitope predictions. *Methods (San Diego, Calif.)* **66**: 3–21.

D'Amato, G, Cecchi, L, Bonini, S, Nunes, C, Annesi-Maesano, I, Behrendt, H, Liccardi, G, Popov, T, van Cauwenberge, P. 2007. Allergenic pollen and pollen allergy in Europe. *Allergy* **62**: 976–990.

Davies, DR, Padlan, EA, Sheriff, S. 1990. Antibody-antigen complexes. *Annual review of biochemistry* **59**: 439–473.

Davies, JM, O'hehir, RE, Suphioglu, C. 2000. Use of phage display technology to investigate allergen-antibody interactions. *The Journal of allergy and clinical immunology* **105**: 1085–1092.

de Monchy, Demoly, P, Akdis, CA, Cardona, V, Papadopoulos, NG, Schmid-Grendelmeier, P, Gayraud, J. 2013. Allergology in Europe, the blueprint. *Allergy* **68**: 1211–1218.

De Swert, L, Gadisseur, R, Sjölander, S, Raes, M, Leus, J, van Hoeyveld, E. 2012. Secondary soy allergy in children with birch pollen allergy may cause both chronic and acute symptoms. *Pediatric allergy and immunology : official publication of the European Society of Pediatric Allergy and Immunology* **23**: 117–123.

Dougherty, DA. 2007. Cation-pi interactions involving aromatic amino acids. *The Journal of nutrition* **137**: 1504S-1508S; discussion 1516S-1517S.

Durham, SR, Emminger, W, Kapp, A, Colombo, G, de Monchy, Jan G R, Rak, S, Scadding, GK, Andersen, JS, Riis, B, Dahl, R. 2010. Long-term clinical efficacy in grass pollen-induced rhinoconjunctivitis after treatment with SQ-standardized grass allergy immunotherapy tablet. *The Journal of allergy and clinical immunology* **125**: 131-8.e1-7.

Durham, SR, Nelson, H. 2011. Allergen immunotherapy: a centenary celebration. *The World Allergy Organization journal* **4**: 104–106.

Ebisawa, M, Brostedt, P, Sjölander, S, Sato, S, Borres, MP, Ito, K. 2013. Gly m 2S albumin is a major allergen with a high diagnostic value in soybean-allergic children. *The Journal of allergy and clinical immunology* **132**: 976-8.e1-5.

Ebo, DG, Bridts, CH, Verweij, MM, De Knop, K J, Hagendorens, MM, De Clerck, L S, Stevens, WJ. 2010. Sensitization profiles in birch pollen-allergic patients with and without oral allergy syndrome to apple: lessons from multiplexed component-resolved allergy diagnosis. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* **40**: 339–347.

Egger, M, Mutschlechner, S, Wopfner, N, Gadermaier, G, Briza, P, Ferreira, F. 2006. Pollen-food syndromes associated with weed pollinosis: an update from the molecular point of view. *Allergy* **61**: 461–476.

El-Hifnawi, D, Neudert, M, Hardt, A, Hauswald, B, Zahnert, T. 2012. Anaphylaktische Reaktion nach Sojagenuss bei einem Patienten mit Birkenpollenallergie. *HNO* **60**: 117–119.

El-Manzalawy, Y, Honavar, V. 2010. Recent advances in B-cell epitope prediction methods. *Immunome research* **6 Suppl 2**: S2.

Ferreira, F, Ebner, C, Kramer, B, Casari, G, Briza, P, Kungl, AJ, Grimm, R, Jahn-Schmid, B, Breiteneder, H, Kraft, D, Breitenbach, M, Rheinberger, HJ, Scheiner, O. 1998. Modulation of IgE reactivity of allergens by site-directed mutagenesis: potential use of hypoallergenic variants for immunotherapy. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **12**: 231–242.

Ferreira, F, Rohlfs, A, Hoffmann-Sommergruber, K, Schenk, S, Ebner, C, Briza, P, Jilek, A, Kraft, D, Breitenbach, M, Scheiner, O. 1996. Modulation of IgE-binding properties of tree pollen allergens by site-directed mutagenesis. *Advances in experimental medicine and biology* **409**: 127–135.

Ferreira, FD, Hoffmann-Sommergruber, K, Breiteneder, H, Pettenburger, K, Ebner, C, Sommergruber, W, Steiner, R, Bohle, B, Sperr, WR, Valent, P. 1993. Purification and characterization of recombinant Bet v I, the major birch pollen allergen. Immunological equivalence to natural Bet v I. *The Journal of biological chemistry* **268**: 19574–19580.

Focke-Tejkl, M, Weber, M, Niespodziana, K, Neubauer, A, Huber, H, Henning, R, Stegfellner, G, Maderegger, B, Hauer, M, Stolz, F, Niederberger, V, Marth, K, Eckl-Dorna, J, Weiss, R, Thalhamer, J, Blatt, K, Valent, P, Valenta, R. 2015. Development and characterization of a recombinant, hypoallergenic, peptide-based vaccine for grass pollen allergy. *The Journal of allergy and clinical immunology* **135**: 1207-7.e1-11.

Franck, P, Moneret Vautrin, D A, Dousset, B, Kanny, G, Nabet, P, Guénard-Bilbaut, L, Parisot, L. 2002. The allergenicity of soybean-based products is modified by food technologies. *International archives of allergy and immunology* **128**: 212–219.

Fraternali, F. 2002. Parameter optimized surfaces (POPS): analysis of key interactions and conformational changes in the ribosome. *Nucleic Acids Research* **30**: 2950–2960.

Fritsch, R, Bohle, B, Vollmann, U, Wiedermann, U, Jahn-Schmid, B, Krebitz, M, Breiteneder, H, Kraft, D, Ebner, C. 1998. Bet v 1, the major birch pollen allergen, and Mal d 1, the major apple allergen, cross-react at the level of allergen-specific T helper cells. *The Journal of allergy and clinical immunology* **102**: 679–686.

Galli, SJ, Tsai, M. 2012. IgE and mast cells in allergic disease. *Nature medicine* **18**: 693–704.

Galli, SJ, Tsai, M, Piliponsky, AM. 2008. The development of allergic inflammation. *Nature* **454**: 445–454.

Ganglberger, E, Grünberger, K, Wiedermann, U, Vermes, M, Sponer, B, Breiteneder, H, Scheiner, O, Boltz, G, Jensen-Jarolim, E. 2001. IgE mimotopes of birch pollen allergen Bet v 1 induce blocking IgG in mice. *International archives of allergy and immunology* **124**: 395–397.

Gehlhar, K, Schlaak, M, Becker, W, Bufe, A. 1999. Monitoring allergen immunotherapy of pollen-allergic patients: the ratio of allergen-specific IgG4 to IgG1 correlates with clinical outcome. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* **29**: 497–506.

Gepp, B, Lengger, N, Bublin, M, Hemmer, W, Breiteneder, H, Radauer, C. 2014. Chimeras of Bet v 1 and Api g 1 reveal heterogeneous IgE responses in patients with birch pollen allergy⋆. *Journal of Allergy and Clinical Immunology*.

Geroldinger-Simic, M, Zelniker, T, Aberer, W, Ebner, C, Egger, C, Greiderer, A, Prem, N, Lidholm, J, Ballmer-Weber, BK, Vieths, S, Bohle, B. 2011. Birch pollen-related food allergy: clinical aspects and the role of allergen-specific IgE and IgG4 antibodies. *The Journal of allergy and clinical immunology* **127**: 616-22.e1.

Gieras, A, Cejka, P, Blatt, K, Focke-Tejkl, M, Linhart, B, Flicker, S, Stoecklinger, A, Marth, K, Drescher, A, Thalhamer, J, Valent, P, Majdic, O, Valenta, R. 2011. Mapping of conformational IgE epitopes with peptide-specific monoclonal antibodies reveals simultaneous binding of different IgE antibodies to a surface patch on the major birch pollen allergen, Bet v 1. *Journal of immunology (Baltimore, Md. : 1950)* **186**: 5333–5344.

González, R, Zapatero, L, Caravaca, F, Carreira, J. 1991. Identification of soybean proteins responsible for respiratory allergies. *International archives of allergy and applied immunology* **95**: 53–57.

Greenfield, N, Fasman, GD. 1969. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **8**: 4108–4116.

Grimshaw, KEC, King, RM, Nordlee, JA, Hefle, SL, Warner, JO, Hourihane, J O B. 2003. Presentation of allergen in different food preparations affects the nature of the allergic reaction--a case series. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* **33**: 1581–1585.

Guex, N, Peitsch, MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**: 2714–2723.

Guhsl, EE, Hofstetter, G, Lengger, N, Hemmer, W, Ebner, C, Fröschl, R, Bublin, M, Lupinek, C, Breiteneder, H, Radauer, C. 2015. IgE, IgG4 and IgA specific to Bet v 1-related food allergens do not predict oral allergy syndrome. *Allergy* **70**: 59–66.

Hansen, KS, Ballmer-Weber, BK, Sastre, J, Lidholm, J, Andersson, K, Oberhofer, H, Lluch-Bernal, M, Ostling, J, Mattsson, L, Schocker, F, Vieths, S, Poulsen, LK. 2009. Component-resolved in vitro diagnosis of hazelnut allergy in Europe. *The Journal of allergy and clinical immunology* **123**: 1134-41, 1141.e1-3.

Haste Andersen, P, Nielsen, M, Lund, O. 2006. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein science : a publication of the Protein Society* **15**: 2558–2567.

Hecker, J, Diethers, A, Schulz, D, Sabri, A, Plum, M, Michel, Y, Mempel, M, Ollert, M, Jakob, T, Blank, S, Braren, I, Spillner, E. 2012. An IgE epitope of Bet v 1 and fagales PR10 proteins as defined by a human monoclonal IgE. *Allergy* **67**: 1530–1537.

Hoffmann-Sommergruber, K. 2000. Plant allergens and pathogenesis-related proteins. What do they have in common? *International archives of allergy and immunology* **122**: 155–166.

Holm, J, Ferreras, M, Ipsen, H, Würtzen, PA, Gajhede, M, Larsen, JN, Lund, K, Spangfort, MD. 2011. Epitope grafting, re-creating a conformational Bet v 1 antibody epitope on the surface of the homologous apple allergen Mal d 1. *The Journal of biological chemistry* **286**: 17569–17578.

Holm, J, Gajhede, M, Ferreras, M, Henriksen, A, Ipsen, H, Larsen, JN, Lund, L, Jacobi, H, Millner, A, Wurtzen, PA, Spangfort, MD. 2004. Allergy Vaccine Engineering: Epitope Modulation of Recombinant Bet v 1 Reduces IgE Binding but Retains Protein Folding Pattern for Induction of Protective Blocking-Antibody Responses. *The Journal of Immunology* **173**: 5258–5267.

Holzhauser, T, Wackermann, O, Ballmer-Weber, BK, Bindslev-Jensen, C, Scibilia, J, Perono-Garoffo, L, Utsumi, S, Poulsen, LK, Vieths, S. 2009. Soybean (Glycine max) allergy in Europe: Gly m 5 (beta-conglycinin) and Gly m 6 (glycinin) are potential diagnostic markers for severe allergic reactions to soy. *The Journal of allergy and clinical immunology* **123**: 452–458.

Holzwarth, G, Doty, P. 1965. The Ultraviolet Circular Dichroism of Polypeptides. *Journal of the American Chemical Society* **87**: 218–228.

Hylander, T, Latif, L, Petersson-Westin, U, Cardell, LO. 2013. Intralymphatic allergen-specific immunotherapy: an effective and safe alternative treatment route for pollen-induced allergic rhinitis. *The Journal of allergy and clinical immunology* **131**: 412–420.

Ilari, A, Franceschini, S, Bonamore, A, Arenghi, F, Botta, B, Macone, A, Pasquo, A, Bellucci, L, Boffi, A. 2009. Structural basis of enzymatic (S)-norcoclaurine biosynthesis. *The Journal of biological chemistry* **284**: 897–904.

Ito, K, Sjölander, S, Sato, S, Movérare, R, Tanaka, A, Söderström, L, Borres, M, Poorafshar, M, Ebisawa, M. 2011. IgE to Gly m 5 and Gly m 6 is associated with severe allergic reactions to soybean in Japanese children. *The Journal of allergy and clinical immunology* **128**: 673–675.

Jahn-Schmid, B, Radakovics, A, Lüttkopf, D, Scheurer, S, Vieths, S, Ebner, C, Bohle, B. 2005. Bet v 1142-156 is the dominant T-cell epitope of the major birch pollen allergen and important for cross-reactivity with Bet v 1-related food allergens. *The Journal of allergy and clinical immunology* **116**: 213–219.

James, LK, Shamji, MH, Walker, SM, Wilson, DR, Wachholz, PA, Francis, JN, Jacobson, MR, Kimber, I, Till, SJ, Durham, SR. 2011. Long-term tolerance after allergen immunotherapy is accompanied by selective persistence of blocking antibodies. *The Journal of allergy and clinical immunology* **127**: 509-516.e1-5.

Jensen-Jarolim, E, Leitner, A, Kalchhauser, H, Zürcher, A, Ganglberger, E, Bohle, B, Scheiner, O, Boltz-Nitulescu, G, Breiteneder, H. 1998. Peptide mimotopes displayed by phage inhibit antibody binding to bet v 1, the major birch pollen allergen, and induce specific IgG response in mice. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **12**: 1635–1642.

Jones, SM, Pons, L, Roberts, JL, Scurlock, AM, Perry, TT, Kulis, M, Shreffler, WG, Steele, P, Henry, KA, Adair, M, Francis, JM, Durham, S, Vickery, BP, Zhong, X, Burks, AW. 2009. Clinical efficacy and immune regulation with peanut oral immunotherapy. *The Journal of allergy and clinical immunology* **124**: 292-300, 300.e1-97.

Jutel, M, Akdis, M, Budak, F, Aebischer-Casaulta, C, Wrzyszcz, M, Blaser, K, Akdis, CA. 2003. IL-10 and TGF-beta cooperate in the regulatory T cell response to mucosal allergens in normal immunity and specific immunotherapy. *European journal of immunology* **33**: 1205–1214.

Jutel, M, Van de Veen, Willem, Agache, I, Azkur, KA, Akdis, M, Akdis, CA. 2013. Mechanisms of allergen-specific immunotherapy and novel ways for vaccine development. *Allergology international : official journal of the Japanese Society of Allergology* **62**: 425–433.

Kahlert, H, Suck, R, Weber, B, Nandy, A, Wald, M, Keller, W, Cromwell, O, Fiebig, H. 2008. Characterization of a hypoallergenic recombinant Bet v 1 variant as a candidate for allergen-specific immunotherapy. *International archives of allergy and immunology* **145**: 193–206.

Karamloo, F, Schmid-Grendelmeier, P, Kussebi, F, Akdis, M, Salagianni, M, von Beust, Barbara R, Reimers, A, Zumkehr, J, Soldatova, L, Housley-Markovic, Z, Müller, U, Kündig, T, Kemeny, DM, Spangfort, MD, Blaser, K, Akdis, CA. 2005. Prevention of allergy by a recombinant multi-allergen vaccine with reduced IgE binding and preserved T cell epitopes. *European journal of immunology* **35**: 3268–3276.

Kay, AB. 2001. Allergy and allergic diseases. First of two parts. *The New England journal of medicine* **344**: 30–37.

Kazemi-Shirazi, L, Pauli, G, Purohit, A, Spitzauer, S, Fröschlc, R, Hoffmann-Sommergruber, K, Breiteneder, H, Scheiner, O, Kraft, D, Valenta, R. 2000. Quantitative IgE inhibition experiments with purified recombinant allergens indicate pollen-derived allergens as the sensitizing agents responsible for many forms of plant food allergy. *Journal of Allergy and Clinical Immunology* **105**: 116–125.

Kelley, LA, Sternberg, MJE. 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols* **4**: 363–371.

Kleine-Tebbe, J, Wangorsch, A, Vogel, L, Crowell, DN, Haustein, U, Vieths, S. 2002. Severe oral allergy syndrome and anaphylactic reactions caused by a Bet v

1– related PR-10 protein in soybean, SAM22. *Journal of Allergy and Clinical Immunology* **110**: 797–804.

Klinglmayr, E, Hauser, M, Zimmermann, F, Dissertori, O, Lackner, P, Wopfner, N, Ferreira, F, Wallner, M. 2009. Identification of B-cell epitopes of Bet v 1 involved in cross-reactivity with food allergens. *Allergy* **64**: 647–651.

Kosma, P, Sjölander, S, Landgren, E, Borres, MP, Hedlin, G. 2011. Severe reactions after the intake of soy drink in birch pollen-allergic children sensitized to Gly m 4. *Acta paediatrica (Oslo, Norway : 1992)* **100**: 305–306.

Laemmli, UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680–685.

Lebecque, ES, Dolecek, C, Laffer, S, Visco, V, Denépoux, S, Pin, JJ, Guret, C, Boltz-Nitulescu, G, Weyer, A, Valenta, R. 1997. Immunologic characterization of monoclonal antibodies that modulate human IgE binding to the major birch pollen allergen Bet v 1. *The Journal of allergy and clinical immunology* **99**: 374–384.

Levin, M, Davies, AM, Liljekvist, M, Carlsson, F, Gould, HJ, Sutton, BJ, Ohlin, M. 2014. Human IgE against the major allergen Bet v 1--defining an epitope with limited cross-reactivity between different PR-10 family proteins. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* **44**: 288–299.

Li, M, Gustchina, A, Alexandratos, J, Wlodawer, A, Wünschmann, S, Kepley, CL, Chapman, MD, Pomés, A. 2008. Crystal structure of a dimerized cockroach allergen Bla g 2 complexed with a monoclonal antibody. *The Journal of biological chemistry* **283**: 22806–22814.

Lins, L, Thomas, A, Brasseur, R. 2003. Analysis of accessible surface of residues in proteins. *Protein science : a publication of the Protein Society* **12**: 1406–1417.

Ma, Y, Gadermaier, G, Bohle, B, Bolhaar, S, Knulst, A, Markovic-Housley, Z, Breiteneder, H, Briza, P, Hoffmann-Sommergruber, K, Ferreira, F. 2006. Mutational analysis of amino acid positions crucial for IgE-binding epitopes of the major apple (Malus domestica) allergen, Mal d 1. *International archives of allergy and immunology* **139**: 53–62.

Mari, A, Ballmer-Weber, BK, Vieths, S. 2005. The oral allergy syndrome: improved diagnostic and treatment methods. *Current opinion in allergy and clinical immunology* **5**: 267–273.

Martínez-Gómez, JM, Johansen, P, Rose, H, Steiner, M, Senti, G, Rhyner, C, Crameri, R, Kündig, TM. 2009. Targeting the MHC class II pathway of antigen presentation enhances immunogenicity and safety of allergen immunotherapy. *Allergy* **64**: 172–178.

Mauro, M, Russello, M, Incorvaia, C, Gazzola, G, Frati, F, Moingeon, P, Passalacqua, G. 2011. Birch-apple syndrome treated with birch pollen immunotherapy. *International archives of allergy and immunology* **156**: 416–422.

May, CD. 1976. Objective clinical and laboratory studies of immediate hypersensitivity reactions to foods in asthmatic children. *The Journal of allergy and clinical immunology* **58**: 500–515.

Mirza, O, Henriksen, A, Ipsen, H, Larsen, JN, Wissenbach, M, Spangfort, MD, Gajhede, M. 2000. Dominant Epitopes and Allergic Cross-Reactivity: Complex

Formation Between a Fab Fragment of a Monoclonal Murine IgG Antibody and the Major Allergen from Birch Pollen Bet v 1. *The Journal of Immunology* **165**: 331–338.

Mittag, D, Akkerdaas, J, Ballmer-Weber, BK, Vogel, L, Wensing, M, Becker, W, Koppelman, SJ, Knulst, AC, Helbling, A, Hefle, SL, van Ree, R, Vieths, S. 2004a. Ara h 8, a Bet v 1-homologous allergen from peanut, is a major allergen in patients with combined birch pollen and peanut allergy. *The Journal of allergy and clinical immunology* **114**: 1410–1417.

Mittag, D, Batori, V, Neudecker, P, Wiche, R, Friis, EP, Ballmer-Weber, BK, Vieths, S, Roggen, EL. 2006. A novel approach for investigation of specific and cross-reactive IgE epitopes on Bet v 1 and homologous food allergens in individual patients. *Molecular immunology* **43**: 268–278.

Mittag, D, Vieths, S, Vogel, L, Becker, W, Rihs, H, Helbling, A, Wüthrich, B, Ballmer-Weber, BK. 2004b. Soybean allergy in patients allergic to birch pollen: Clinical investigation and molecular characterization of allergens. *Journal of Allergy and Clinical Immunology* **113**: 148–154.

Moldaver, D, Larché, M. 2011. Immunotherapy with peptides. *Allergy* **66**: 784–791.

Mueller, GA, Smith, AM, Chapman, MD, Rule, GS, Benjamin, DC. 2001. Hydrogen exchange nuclear magnetic resonance spectroscopy mapping of antibody epitopes on the house dust mite allergen Der p 2. *The Journal of biological chemistry* **276**: 9359–9365.

Murphy, K, Travers, P, Walport, M, Janeway, C. 2012. Janeway's immunobiology. New York: Garland Science.

Naik, MT, Chang, C, Kuo, I, Kung, CC, Yi, F, Chua, K, Huang, T. 2008. Roles of structure and structural dynamics in the antibody recognition of the allergen proteins: an NMR study on Blomia tropicalis major allergen. *Structure (London, England : 1993)* **16**: 125–136.

Negi, SS, Braun, W. 2009. Automated detection of conformational epitopes using phage display Peptide sequences. *Bioinformatics and biology insights* **3**: 71–81.

Neudecker, P, Lehmann, K, Nerkamp, J, Haase, T, Wangorsch, A, Fötisch, K, Hoffmann, S, Rösch, P, Vieths, S, Scheurer, S. 2003a. Mutational epitope analysis of Pru av 1 and Api g 1, the major allergens of cherry (Prunus avium) and celery (Apium graveolens): correlating IgE reactivity with three-dimensional structure. *The Biochemical journal* **376**: 97–107.

Neudecker, P, Lehmann, K, Rösch, P. 2003b. Sequence-specific 1H, 13C and 15N resonance assignments of SAM22, an allergenic stress-induced protein from soy bean. *Journal of biomolecular NMR* **26**: 191–192.

Neudecker, P, Schweimer, K, Nerkamp, J, Scheurer, S, Vieths, S, Sticht, H, Rösch, P. 2001. Allergic cross-reactivity made visible: solution structure of the major cherry allergen Pru av 1. *The Journal of biological chemistry* **276**: 22756–22763.

Niemi, M, Jylhä, S, Laukkanen, M, Söderlund, H, Mäkinen-Kiljunen, S, Kallio, JM, Hakulinen, N, Haahtela, T, Takkinen, K, Rouvinen, J. 2007. Molecular interactions between a recombinant IgE antibody and the beta-lactoglobulin allergen. *Structure (London, England : 1993)* **15**: 1413–1421.

Nwaru, BI, Hickstein, L, Panesar, SS, Roberts, G, Muraro, A, Sheikh, A. 2014. Prevalence of common food allergies in Europe: a systematic review and meta-analysis. *Allergy* **69**: 992–1007.

Padavattan, S, Flicker, S, Schirmer, T, Madritsch, C, Randow, S, Reese, G, Vieths, S, Lupinek, C, Ebner, C, Valenta, R, Markovic-Housley, Z. 2009. High-affinity IgE recognition of a conformational epitope of the major respiratory allergen Phl p 2 as revealed by X-ray crystallography. *Journal of immunology (Baltimore, Md. : 1950)* **182**: 2141–2151.

Padavattan, S, Schirmer, T, Schmidt, M, Akdis, C, Valenta, R, Mittermann, I, Soldatova, L, Slater, J, Mueller, U, Markovic-Housley, Z. 2007. Identification of a B-cell epitope of hyaluronidase, a major bee venom allergen, from its crystal structure in complex with a specific Fab. *Journal of molecular biology* **368**: 742–752.

Page, R, Peti, W, Wilson, IA, Stevens, RC, Wüthrich, K. 2005. NMR screening and crystal quality of bacterially expressed prokaryotic and eukaryotic proteins in a structural genomics pipeline. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 1901–1905.

Piotto, M, Saudek, V, Sklenár, V. 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *Journal of biomolecular NMR* **2**: 661–665.

Pomés, A. 2010. Relevant B cell epitopes in allergic disease. *International archives of allergy and immunology* **152**: 1–11.

Ponomarenko, J, Bui, H, Li, W, Fusseder, N, Bourne, PE, Sette, A, Peters, B. 2008. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC bioinformatics* **9**: 514.

Poulsen, LK, Hummelshoj, L. 2007. Triggers of IgE class switching and allergy development. *Annals of medicine* **39**: 440–456.

Pree, I, Reisinger, J, Focke, M, Vrtala, S, Pauli, G, van Hage, M, Cromwell, O, Gadermaier, E, Egger, C, Reider, N, Horak, F, Valenta, R, Niederberger, V. 2007. Analysis of Epitope-Specific Immune Responses Induced by Vaccination with Structurally Folded and Unfolded Recombinant Bet v 1 Allergen Derivatives in Man. *The Journal of Immunology* **179**: 5309–5316.

Radauer, C, Lackner, P, Breiteneder, H. 2008. The Bet v 1 fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands. *BMC evolutionary biology* **8**: 286.

Rajan, TV. 2003. The Gell–Coombs classification of hypersensitivity reactions: a re-interpretation. *Trends in Immunology* **24**: 376–379.

Razzera, G, Gadermaier, G, Paula, V de, Almeida, MS, Egger, M, Jahn-Schmid, B, Almeida, Fabio C L, Ferreira, F, Valente, AP. 2010. Mapping the interactions between a major pollen allergen and human IgE antibodies. *Structure (London, England : 1993)* **18**: 1011–1021.

Reese, G, Ayuso, R, Leong-Kee, SM, Plante, MJ, Lehrer, SB. 2001. Characterization and identification of allergen epitopes: recombinant peptide libraries and synthetic, overlapping peptides. *Journal of chromatography. B, Biomedical sciences and applications* **756**: 157–163.

Reuter, A, Lidholm, J, Andersson, K, Ostling, J, Lundberg, M, Scheurer, S, Enrique, E, Cistero-Bahima, A, San Miguel-Moncin, M, Ballmer-Weber, BK, Vieths, S. 2006. A critical assessment of allergen component-based in vitro diagnosis in cherry allergy across Europe. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* **36**: 815–823.

Rihs, HP, Chen, Z, Ruëff, F, Petersen, A, Rozynek, P, Heimann, H, Baur, X. 1999. IgE binding of the recombinant allergen soybean profilin (rGly m 3) is mediated by conformational epitopes. *The Journal of allergy and clinical immunology* **104**: 1293–1301.

Robert, N, Ferran, J, Breda, C, Coutos-Thévenot, P, Boulay, M, Buffard, D, Esnault, R. 2001. Molecular Characterization of the Incompatible Interaction of Vitis Vinifera Leaves With Pseudomonas Syringae Pv. Pisi: Expression Of Genes Coding For Stilbene Synthase And Class 10 PR Protein. *European Journal of Plant Pathology* **107**: 249–261. http://download.springer.com/static/pdf/180/art%253A10.1023%252FA%253A1011241001383.pdf?auth66=1411214831_7624ceaa127ec2de76d424648756e972&ext=.pdf.

Romagnani, S. 2004. Immunologic influences on allergy and the TH1/TH2 balance. *The Journal of allergy and clinical immunology* **113**: 395–400.

Rougé, P, Culerrier, R, Sabatier, V, Granier, C, Rancé, F, Barre, A. 2009. Mapping and conformational analysis of IgE-binding epitopic regions on the molecular surface of the major Ara h 3 legumin allergen of peanut (Arachis hypogaea). *Molecular Immunology* **46**: 1067–1075.

Samanani, N, Liscombe, DK, Facchini, PJ. 2004. Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis. *The Plant journal : for cell and molecular biology* **40**: 302–313.

Sampson, HA. 2001. Utility of food-specific IgE concentrations in predicting symptomatic food allergy. *The Journal of allergy and clinical immunology* **107**: 891–896.

Sampson, HA, Ho, DG. 1997. Relationship between food-specific IgE concentrations and the risk of positive food challenges in children and adolescents. *The Journal of allergy and clinical immunology* **100**: 444–451.

Saraste, M, Sibbald, PR, Wittinghofer, A. 1990. The P-loop — a common motif in ATP- and GTP-binding proteins. *Trends in Biochemical Sciences* **15**: 430–434.

Saxon, A, Diaz-Sanchez, D. 2005. Air pollution and allergy: you are what you breathe. *Nature immunology* **6**: 223–226.

Scheurer, S, Son, D, Boehm, M, Karamloo, F, Franke, S, Hoffmann, A, Haustein, D, Vieths, S. 1999. Cross-reactivity and epitope analysis of Pru a 1, the major cherry allergen. *Molecular Immunology* **36**: 155–167.

Schimek, EM, Zwölfer, B, Briza, P, Jahn-Schmid, B, Vogel, L, Vieths, S, Ebner, C, Bohle, B. 2005. Gastrointestinal digestion of Bet v 1-homologous food allergens destroys their mediator-releasing, but not T cell-activating, capacity. *The Journal of allergy and clinical immunology* **116**: 1327–1333.

Schülke, S, Wolfheimer, S, Gadermaier, G, Wangorsch, A, Siebeneicher, S, Briza, P, Spreitzer, I, Schiller, D, Loeschner, B, Uematsu, S, Ryffel, B, Akira, S, Waibler, Z, Vieths, S, Toda, M, Scheurer, S. 2014. Prevention of intestinal allergy in mice by rflaA:Ova is associated with enforced antigen processing and TLR5-dependent IL-10 secretion by mDC. *PLoS ONE* **9**: e87822.

Senti, G, Moos, S von, Tay, F, Graf, N, Sonderegger, T, Johansen, P, Kündig, TM. 2012. Epicutaneous allergen-specific immunotherapy ameliorates grass pollen-induced rhinoconjunctivitis: A double-blind, placebo-controlled dose escalation study. *The Journal of allergy and clinical immunology* **129**: 128–135.

Seutter von Loetzen, C, Hoffmann, T, Hartl, MJ, Schweimer, K, Schwab, W, Rösch, P, Hartl-Spiegelhauer, O. 2014. Secret of the major birch pollen allergen Bet v 1: identification of the physiological ligand. *The Biochemical journal* **457**: 379–390.

Shreffler, WG, Lencer, DA, Bardina, L, Sampson, HA. 2005. IgE and IgG4 epitope mapping by microarray immunoassay reveals the diversity of immune response to the peanut allergen, Ara h 2. *The Journal of allergy and clinical immunology* **116**: 893–899.

Sicherer, SH. 2011. Epidemiology of food allergy. *The Journal of allergy and clinical immunology* **127**: 594–602.

Sicherer, SH, Sampson, HA. 2014. Food allergy: Epidemiology, pathogenesis, diagnosis, and treatment. *The Journal of allergy and clinical immunology* **133**: 291-307; quiz 308.

Sicherer, SH, Teuber, S. 2004. Current approach to the diagnosis and management of adverse reactions to foods. *The Journal of allergy and clinical immunology* **114**: 1146–1150.

Silva, JC, Denny, R, Dorschel, CA, Gorenstein, M, Kass, IJ, Li, G, McKenna, T, Nold, MJ, Richardson, K, Young, P, Geromanos, S. 2005. Quantitative proteomic analysis by accurate mass retention time pairs. *Analytical chemistry* **77**: 2187–2200.

Silva, JC, Gorenstein, MV, Li, G, Vissers, Johannes P C, Geromanos, SJ. 2006. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Molecular & cellular proteomics : MCP* **5**: 144–156.

Somssich, IE, Schmelzer, E, Bollmann, J, Hahlbrock, K. 1986. Rapid activation by fungal elicitor of genes encoding "pathogenesis-related" proteins in cultured parsley cells. *Proceedings of the National Academy of Sciences of the United States of America* **83**: 2427–2430.

Son, DY, Scheurer, S, Hoffmann, A, Haustein, D, Vieths, S. 1999. Pollen-related food allergy: cloning and immunological analysis of isoforms and mutants of Mal d 1, the major apple allergen, and Bet v 1, the major birch pollen allergen. *European journal of nutrition* **38**: 201–215.

Spangfort, MD, Mirza, O, Ipsen, H, van Neerven, R. J. J., Gajhede, M, Larsen, JN. 2003. Dominating IgE-Binding Epitope of Bet v 1, the Major Allergen of Birch Pollen, Characterized by X-ray Crystallography and Site-Directed Mutagenesis. *The Journal of Immunology* **171**: 3084–3090.

Spangfort, MD, Mirza, O, Svensson, LA, Larsen, JN, Gajhede, M, Ipsen, H. 1999. Crystallization and preliminary X-ray analysis of birch-pollen allergen Bet v 1 in complex with a murine monoclonal IgG Fab′ fragment. *Acta Crystallographica Section D Biological Crystallography* **55**: 2035–2036.

Spergel, JM, Beausoleil, JL, Fiedler, JM, Ginsberg, J, Wagner, K, Pawlowski, NA. 2004. Correlation of initial food reactions to observed reactions on challenges. *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology* **92**: 217–224.

Steckelbroeck, S, Ballmer-Weber, BK, Vieths, S. 2008. Potential, pitfalls, and prospects of food allergy diagnostics with recombinant allergens or synthetic sequential epitopes. *The Journal of allergy and clinical immunology* **121**: 1323–1330.

Strachan, DP. 1989. Hay fever, hygiene, and household size. *BMJ (Clinical research ed.)* **299**: 1259–1260.

Subbarayal, B, Schiller, D, Möbs, C, de Jong, N W, Ebner, C, Reider, N, Bartel, D, Lidholm, J, Pfützner, W, Gerth van Wijk, R, Vieths, S, Bohle, B. 2013. Kinetics, cross-reactivity, and specificity of Bet v 1-specific IgG4 antibodies induced by immunotherapy with birch pollen. *Allergy* **68**: 1377–1386.

Syed, A, Kohli, A, Nadeau, KC. 2013. Food allergy diagnosis and therapy: where are we now? *Immunotherapy* **5**: 931–944.

Tolkki, L, Alanko, K, Petman, L, Skydtsgaard, MB, Milvang, PG, Seppälä, U, Ranki, A. 2013. Clinical characterization and IgE profiling of birch (Betula verrucosa)--allergic individuals suffering from allergic reactions to raw fruits and vegetables. *The journal of allergy and clinical immunology. In practice* **1**: 623-31.e1.

Tomicić, S, Norrman, G, Fälth-Magnusson, K, Jenmalm, MC, Devenney, I, Böttcher, MF. 2009. High levels of IgG4 antibodies to foods during infancy are associated with tolerance to corresponding foods later in life. *Pediatric allergy and immunology : official publication of the European Society of Pediatric Allergy and Immunology* **20**: 35–41.

Tordesillas, L, Pacios, LF, Palacín, A, Cuesta-Herranz, J, Madero, M, Díaz-Perales, A. 2010. Characterization of IgE epitopes of Cuc m 2, the major melon allergen, and their role in cross-reactivity with pollen profilins. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* **40**: 174–181.

van Ree, R, Van Leeuwen, W A, Dieges, PH, Van Wijk, R G, Jong, N de, Brewczyski, PZ, Kroon, AM, Schilte, PP, Tan, KY, Simon-Licht, IF, Roberts, AM, Stapel, SO, Aalberse, RC. 1997. Measurement of IgE antibodies against purified grass pollen allergens (Lol p 1, 2, 3 and 5) during immunotherapy. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* **27**: 68–74.

Van Regenmortel, MH. 1989. Structural and functional approaches to the study of protein antigenicity. *Immunology today* **10**: 266–272.

Vázquez-Ortiz, M, Alvaro-Lozano, M, Alsina, L, Garcia-Paba, MB, Piquer-Gibert, M, Giner-Muñoz, MT, Lozano, J, Domínguez-Sánchez, O, Jiménez, R, Días, M,

Martín-Mateos, MA, Plaza-Martín, AM. 2013. Safety and predictors of adverse events during oral immunotherapy for milk allergy: severity of reaction at oral challenge, specific IgE and prick test. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* **43**: 92–102.

Venyaminov, Baikalov, IA, Shen, ZM, Wu, CS, Yang, JT. 1993. Circular dichroic analysis of denatured proteins: inclusion of denatured proteins in the reference set. *Analytical Biochemistry* **214**: 17–24.

Vercelli, D. 2008. Discovering susceptibility genes for asthma and allergy. *Nature reviews. Immunology* **8**: 169–182.

Vereda, A, Andreae, DA, Lin, J, Shreffler, WG, Ibañez, MD, Cuesta-Herranz, J, Bardina, L, Sampson, HA. 2010. Identification of IgE sequential epitopes of lentil (Len c 1) by means of peptide microarray immunoassay. *The Journal of allergy and clinical immunology* **126**: 596-601.e1.

Vieths, S, Scheurer, S, Ballmer-Weber, BK. 2002. Current Understanding of Cross-Reactivity of Food Allergens and Pollen. *Annals of the New York Academy of Sciences* **964**: 47–68.

Vogel, L, Lüttkopf, D, Hatahet, L, Haustein, D, Vieths, S. 2005. Development of a functional in vitro assay as a novel tool for the standardization of allergen extracts in the human system. *Allergy* **60**: 1021–1028.

Wallner, M, Hauser, M, Himly, M, Zaborsky, N, Mutschlechner, S, Harrer, A, Asam, C, Pichler, U, van Ree, R, Briza, P, Thalhamer, J, Bohle, B, Achatz, G, Ferreira, F. 2011. Reshaping the Bet v 1 fold modulates T(H) polarization. *The Journal of allergy and clinical immunology* **127**: 1571-8.e9.

Walter, MH, Liu, JW, Wünn, J, Hess, D. 1996. Bean ribonuclease-like pathogenesis-related protein genes (Ypr10) display complex patterns of developmental, dark-induced and exogenous-stimulus-dependent expression. *European journal of biochemistry / FEBS* **239**: 281–293.

Wangorsch, A, Ballmer-Weber, BK, Rösch, P, Holzhauser, T, Vieths, S. 2007. Mutational epitope analysis and cross-reactivity of two isoforms of Api g 1, the major celery allergen. *Molecular immunology* **44**: 2518–2527.

Warner, SA, Scott, R, Draper, J. 1992. Characterisation of a wound-induced transcript from the monocot asparagus that shares similarity with a class of intracellular pathogenesis-related (PR) proteins. *Plant molecular biology* **19**: 555–561.

White, MV. 1990. The role of histamine in allergic diseases. *The Journal of allergy and clinical immunology* **86**: 599–605.

Wiche, R, Gubesch, M, König, H, Fötisch, K, Hoffmann, A, Wangorsch, A, Scheurer, S, Vieths, S. 2005. Molecular basis of pollen-related food allergy: identification of a second cross-reactive IgE epitope on Pru av 1, the major cherry (Prunus avium) allergen. *The Biochemical journal* **385**: 319–327.

Wills-Karp, M, Santeliz, J, Karp, CL. 2001. The germless theory of allergic disease: revisiting the hygiene hypothesis. *Nature reviews. Immunology* **1**: 69–75.

8 APPENDIX

8.1 Supporting information

Table A 1: Peptide and mimotope data resulting from Mittag *et al.* using phage-display and competitive immunoscreening with rGly m 4 (Mittag *et al.*, 2006).

Peptide	Mimotope	Center residue	Score	Length
LEKGTPW	G1, T4, L90, T93, E95, K96, T98, K118, E120, T121, E126, P127, L132; G1, G86, G87, L90, P91, T93, E95, K96, E120, T121, K122, G123, P127;	119 94	1	13
AAKGTPW	G1, G86, G87, A88, A89, P91, T93, A94, K96, T121, K122, G123, A125, P127;	94	1	14
QVELPKG	E41, V43, K53, K54, V66, L67, K69, V84, V85, G86, G87, L90, P91;	67	1	13
KLPLTKG	G1, G86, G87, L90, P91, T93, K96, T121, K122, G123, P127;	94	1	11
QKTAASY	S72, Y80, S81, Y82, T98, S101, K102, K114, T116, K118 ;	100	1	10
EAAKQGE	G1, G86, G87, A88, A89, A94, E95, K96, E120, K122, G123, A125;	94	1	14
IKLPALF	A26, I30, P31, K32, A33, L34, F37, K38, I55, F57, L58, A141, L142, A145;	37	1	14
KNLEEPK	L90, E95, K96, K118, E120, E126, P127, L132; K20, L22, N28, P31, K32, L34, L142, L150; N10, P12, L22, L142, K144, E147, L150, L151; N10, L22, L115, K136, K138, L142, K144, E147;	119 29 147 143	1	8
KYLEEPV	V2, V84, L90, E95, K96, V117, K118, Y119, E120, E126, P127, L132; V66, Y82, V84, V85, L90, E95, K96, V117, K118, Y119, E120, P127; V2, K69, E71, Y82, V84, V85, E95, K96, V117, K118, Y119, E120;	119 97 98	1	12
TPIRHRS	P49, H68, S72, I73, S81, I97, T98, S101;	82	1	8
QYPWRLA	P12, A14, A21, L22, A145, A148, Y149, L150, L151, A152, P154, Y156;	150	1	12
FDMATNP	F3, F5, A89. T93, A94, T121, A125, P127, N128, D130; F3, T4, F5, T93, A94, T98, F99, T121, A125, P127; A21, T24, D25, A26, D27, N28, P31, A33, F37, A145; F3, A88, A89, P91, D92, T93, A94, T121, A125, P127; A26, D27, P31, A33, D35, F37, T56, F57, A141, A145;	127 119 129 94 37	1	10
SLYELTH	F3, T4, F5, D7, T98, F99, D100, T116, P127, A139 ; L22, T24, E71, S72, E75, L78, Y80, S81, Y82, S101; T4, L90, T93, E95, T98, Y119, E120, T121, E126, L132; H68, E71, Y80, S81, Y82, T98, S101, L115, T116, Y119; E6, E71, S72, Y80, S81, Y82, T98, S101, L115, T116;	117 80 119 99 100	1	10

TAIHAVW	V66, H68, V84, V85, A88, A89, A94, I97, T98, V117, T121;	97	1	11
DHLPSPW	L22, D25, D27, P31, L34, S36, S39, L142; D27, P31, L34, D35, S36, S39, L58, L142;	30 37	1	8
YKSDNIR	S72, I73, D74, N77, Y80, S81, Y82, D100, S101, K102; K53, K69, S72, I73, Y80, S81, Y82, I97, D100, S101; D25, D27, N28, I30, K32, S39, N42, K53, K54, I55; S72, I73, D74, N77, Y80, S81, D100, S101, K102, K114;	80 82 27 79	1	10
SKTDNLT	T4, D7, S81, T98, D100, S101, K102, K114, L115, T116, K118;	116	1	11
AFNRASD	A21, D25, A26, D27, N28, A33, S36, F37, S39, A145; A26, D27, A33, D35, S36, F37, S39, F57, A141, A145;	30 37	1	10
YLMEVQK	V2, K69, E71, Y82, V84, V85, E95, K96, V117, K118, Y119, E120;	98	1	12
YEMDAPK	A88, A89, P91, D92, A94, E95, K96, Y119, E120, K122, A125, P127;	94	1	12
ERHLTFT	F3, T4, F5, L90, T93, E95, T98, F99, E120, T121, E126, L132;	119	1	12

21 peptides for soybean allergic patients were identified and mapped onto Gly m 4 surface using EpiSearch resulting in 40 potential IgE-binding mimotopes. Resulting from EpiSearch analysis, center residue, score and total length of each identified mimotope are listed.

Table A 2: Peptide and mimotope data from bioinformatical approach using 7mer peptides representing total amino acid sequence of Gly m 4.

Peptide	Mimotope	Center residue	Score	Length
MGVFTFE	G1, V2, F3, T4, F5, V84, T93, E95, T98, F99, V117, E120, T121, E126;	119	1	14
GVFTFED	G1, V2, F3, T4, F5, V84, T93, E95, T98, F99, V117, E120, T121, E126;	119	1	14
VFTFEDE	V2, F3, T4, F5, V84, T93, E95, T98, F99, V117, E120, T121, E126;	119	1	13
FTFEDEI	F3, T4, F5, T93, E95, I97, T98, F99, E120, T121, E126;	119	1	11
TFEDEIN	F3, T4, F5, T93, E95, I97, T98, F99, E120, T121, E126;	119	1	11
FEDEINS	E71, S72, I73, D74, E75, N77, S81, F99, D100, S101; D35, S36, F37, S39, I55, F57, E59, D60, E62, F65;	80 57	1	10
EDEINSP	E71, S72, I73, D74, E75, N77, S81, D100, S101; D25, D27, N28, I30, P31, S39, E41, N42, I55;	80 27	1	9
DEINSPV	D25, D27, N28, V29, I30, P31, S39, V40, E41, N42, I55; D27, S39, V40, E41, N42, V43, E44, I52, I55, V66, V85;	27 54	1	11
EINSPVA	A21, A26, N28, V29, I30, P31, A33, S36, S39, V40, I55, A145;	30	1	12

INSPVAP	A21, A26, N28, V29, I30, P31, A33, S36, S39, V40, I55, A145;	30	1	12
NSPVAPA	A21, A26, N28, V29, P31, A33, S36, S39, V40, A145;	30	1	10
SPVAPAT	V2, T4, V84, T93, A94, T98, V117, T121, A125, P127; P12, V13, A14, T17, A21, V29, A145, A148, A152, P154; V66, V84 V85, A88, A89, A94, T98, V117, T121, P127; A26, V29, P31, A33, S36, S39, V40, T56, A141, A145;	119 150 97 37	1	10
PVAPATL	P12, V13, A14, T17, A21, L22, V29, A145, A148, L150, L151, A152, P154;	150	1	13
VAPATLY	P12, V13, A14, T17, A21, L22, V29, A145, A148, Y149, L150, L151, A152, P154, Y156;	150	1	15
APATLYK	P12, A14, T17, A21, L22, A145, A148, Y149, L150, L151, A152, P154, Y156; K20, A21, L22, T24, A26, P31, K32, A33, L34, L142, A145, Y149, L150;	150 29	1	13
PATLYKA	P12, A14, T17, A21, L22, A145, A148, Y149, L150, L151, A152, P154, Y156; K20, A21, L22, T24, A26, P31, K32, A33, L34, L142, A145, Y149, L150;	150 29	1	13
ATLYKAL	K20, A21, L22, T24, A26, K32, A33, L34, L142, A145, Y149, L150; A21, K32, A33, L34, K144, A145, A148, Y149, L150, L151, A152, Y156;	29 149	1	12
TLYKALV	V2, T4, V84, L90, T93, A94, K96, T98, V117, K118, Y119, T121, A125, L132; K20, A21, L22, T24, A26, V29, K32, A33, L34, V40, L142, A145, Y149, L150; V13, A21, V29, K32, A33, L34, K144, A145, A148, Y149, L150, L151, A152, Y156; V66, Y82, V84, V85, A88, A89, L90, A94, K96, T98, V117, K118, Y119, T121,	119 29 149 97	1	14
LYKALVT	V2, T4, V84, L90, T93, A94, K96, T98, V117, K118, Y119, T121, A125, L132; K20, A21, L22, T24, A26, V29, K32, A33, L34, V40, L142, A145, Y149, L150; V13, A21, V29, K32, A33, L34, K144, A145, A148, Y149, L150, L151, A152, Y156; V66, Y82, V84, V85, A88, A89, L90, A94, K96, T98, V117, K118, Y119, T121;	119 29 149 97	1	14
YKALVTD	K20, A21, L22, T24, D25, A26, D27, V29, K32, A33, L34, V40, L142, A145, Y149, L150;	29		16
KALVTDA	K20, A21, L22, T24, D25, A26, D27, V29, K32, A33, L34, V40, L142, A145, L150;	29	1	15
ALVTDAD	A21, L22, T24, D25, A26, D27, V29, A33, L34, V40, L142, A145, L150;	29	1	13
LVTDADN	A21, L22, T24, D25, A26, D27, N28, V29, A33, L34, V40, L142, A145, L150;	29	1	14
VTDADNV	A21, T24, D25, A26, D27, N28, V29, A33, V40, A145;	29	1	10
TDADNVI	A21, D25, A26, D27, N28, V29, I30, A33, V40, I55, A145; A21, T24, D25, A26, D27, N28, V29, I30, A33, V40, A145; A26, D27, V29, I30, A33, D35, V40, I55, T56, A141, A145;	30 29 37	1	11
DADNVIP	A21, D25, A26, D27, N28, V29, I30, P31, A3 , V40, I55, A145;	30	1	12
ADNVIPK	A26, D27, V40, V43, I52, K53, K54, I55, V66, K69, V84, V85, I97; A21, D25, A26, D27, N28, V29, I30, P31, K32, A33, V40, I55, A145; D25, A26, D27, N28, V29, I30, P31, K32, V40, N42, K53, K54, I55; K20, A21, D25, A26, D27, N28, V29, I30, P31, K32, A33, V40, A145; A26, D27, V29, I30, P31, K32, A33, D35, K38, V40, I55, A141, A145;	68 30 27 29 37	1	13

DNVIPKA	A26, D27, V40, V43, I52, K53, K54, I55, V66, K69, V84, V85, I97; A21, D25, A26, D27, N28, V29, I30, P31, K32, A33, V40, I55, A145; D25, A26, D27, N28, V29, I30, P31, K32, V40, N42, K53, K54, I55; K20, A21, D25, A26, D27, N28, V29, I30, P31, K32, A33, V40, A145; A26, D27, V29, I30, P31, K32, A33, D35, K38, V40, I55, A141, A145;	68 30 27 29 37	1	13
NVIPKAL	K20, A21, L22, A26, N28, V29, I30, P31, K32, A33, L34, V40, L142, A145, L150;	29	1	15
VIPKALD	K20, A21, L22, D25, A26, D27, V29, I30, P31, K32, A33, L34, V40, L142, A145, L150; A26, D27, V29, I30, P31, K32, A33, L34, D35, K38, V40, I55, L58, A141, L142, A145;	29 37	1	16
IPKALDS	A26, D27, I30, P31, K32, A33, L34, D35, S36, K38, S39, I55, L58, A141, L142, A145;	37	1	16
PKALDSF	A26, D27, P31, K32, A33, L34, D35, S36, F37, K38, S39, F57, L58, A141, L142, A145;	37	1	16
KALDSFK	A26, D27, K32, A33, L34, D35, S36, F37, K38, S39, F57, L58, A141, L142, A145;	37	1	15
ALDSFKS	A21, L22, D25, A26, D27, K32, A33, L34, S36, F37, S39, L142, A145; K20, A21, L22, D25, A26, D27, K32, A33, L34, F37, L142, A145, L150;	30 29	1	13
LDSFKSV	D27, V29, K32, L34, D35, S36, F37, K38, S39, V40, F57, L58, L142;	37	1	13
DSFKSVE	D35, S36, F37, K38, S39 F57, E59, D60, E62, K64, F65, V66, K138; V2, K69, E71, S81, V84, V85, E95, K96, F99, D100, V117, K118, E120;	57 98	1	13
SFKSVEN	V2, K69, E71, S81, V84, V85, E95, K96, F99, V117, K118, E120; S39, V40, E41, N42, V43, E44, K53, K54, F65, V66, K69, V85;	98 54	1	12
FKSVENV	V2, K69, E71, S81, V84, V85, E95, K96, F99, V117, K118, E120; S39, V40, E41, N42, V43, E44, K53, K54, F65, V66, K69, V85;	98 54	1	12
KSVENVE	V2, K69, E71, S81, V84, V85, E95, K96, V117, K118, E120; S39, V40, E41, N42, V43, E44, K53, K54, V66, K69, V85;	98 54	1	11
SVENVEG	S39, V40, E41, N42, V43, E44, V66, V85, G86, G87;	54	1	10
VENVEGN	V40, E41, N42, V43, E44, G45, G47, G50, V66; V40, E41, N42, V43, E44, V66, V85, G86, G87;	53 54	1	9
ENVEGNG	V40, E41, N42, V43, E44, G45, G47, G50, V66; V40, E41, N42, V43, E44, V66, V85, G86, G87;	53 54	1	9
NVEGNGG	V40, E41, N42, V43, E44, G45, G47, G50, V66 V40, E41, N42, V43 , E44, V66, V85, G86, G87;	53 54	1	9
VEGNGGP	V40, E41, N42, V43, E44, G45, G47, G50, V66; G1, V84, G86, G87, P91, E95, E120, G123, P127; N42, V43, E44, G45, N46, G47, G48, P49, G50; V40, E41, N42, V43, E44, V66, V85, G86, G87;	53 94 45 45	1	9
EGNGGPG	E44, G45, N46, G47, G48, P49, G50, E71; G1, G86, G87, P91, E95, E120, G123, P127; N42, E44, G45, N46, G47, G48, P49, G50; N42, E44, G45, N46, G47, G48, P49, G50;	50 94 47 45	1	8
GNGGPGT	G1, G86, G87, P91, T93, T121, G123, P127; T24, N42, G45, N46, G47, G48, P49, G50;	94 47	1	8

NGGPGTI	T24, N42, G45, N46, G47, G48, P49, G50, I52, I73;	47	1	10
GGPGTIK	G1, G86, G87, P91, T93, K96, I97, T121, K122, G123, P127;	94	1	11
GPGTIKK	G1, G86, G87, P91, T93, K96, I97, T121, K122, G123, P127;	94	1	11
PGTIKKI	G1, G86, G87, P91, T93, K96, I97, T121, K122, G123, P127;	94	1	11
GTIKKIT	G1, G86, T93, K96, I97, T98, T121, K122, G123; K38, K54, I55, T56, G61, T63 K64, G86, G87; K53, K54, I55, T56, T63, K64, G86, G87, I97;	95 65 66	1	9
TIKKITF	G1, G86, G87, T93, K96, I97, T121, K122, G123; F3, T4, F5, T93, K96, I97, T98, F99, K118, T121; I30, F37, K38, K53, K54, I55, T56, F57, T63, F65;	94 119 55	1	10
IKKITFL	F3, T4, F5, L90, T93, K96, I97, T98, F99, K118, T121, L132;	119	1	12
KKITFLE	F3, T4, F5, L90, T93, E95, K96, I97, T98, F99, K118, E120, T121, E126, L132;	119	1	15
KITFLED	F3, T4, F5, L90, T93, E95, K96, I97, T98, F99, K118, E120, T121, E126, L132;	119	1	15
ITFLEDG	G1, F3, T4, F5, L90, T93, E95, I97, T98, F99, E120, T121, E126, L132;	119	1	14
TFLEDGE	G1, F3, T4, F5, L90, T93, E95, T98, F99, E120, T121, E126, L132;	119	1	13
FLEDGET	G1, F3, T4, F5, L90, T93, E95, T98, F99, E120, T121, E126, L132;	119	1	13
LEDGETK	G1, T4, L90, T93, E95, K96, T98, K118, E120, T121, E126, L132; G1, G86, G87, L90, D92, T93, E95, K96, E120, T121, K122, G123;	119 94	1	12
EDGETKF	G1, F3, T4, F5, T93, E95, K96, T98, F99, K118, E120, T121, E126; D35, F37, K38, T56, F57, E59, D60, G61, E62, T63, K64, F65, K138;	119 57	1	13
DGETKFV	G1, V2, F3, T4, F5, V84, T93, E95, K96, T98, F99, V117, K118, E120, T121, E126;	119	1	16
GETKFVL	G1, V2, F3, T4, F5, V84, L90, T93, E95, K96, T98, F99, V117, K118, E120, T121, E126, L132;	119	1	18
ETKFVLH	V2, F3, T4, F5, V84, L90, T93, E95, K96, T98, F99, V117, K118, E120, T121, E126, L132;	119	1	17
TKFVLHK	V2, F3, T4, F5, V84, L90, T93, K96, T98, F99, V117, K118, T121, L132;	119	1	17
KFVLHKI	V40, V43, I52, K53, K54, I55, F65, V66, L67, H68, K69, V84, V85, I97, F99;	68	1	15
FVLHKIE	V40, V43, I52, K53, K54, I55, F65, V66, L67, H68, K69, E71, V84, V85, I97, F99;	68	1	16
VLHKIES	V40, V43, I52, K53, K54, I55, V66, L67, H68, K69, E71, V84, V85, I97; S39, V40, E41, V43, E44, I52, K53, K54, I55, V66, L67, H68, K69, V85;	68 54	1	14
LHKIESI	S39, E41, E44, I52, K53, K54, I55, L67, H68, K69;	54	1	10
HKIESID	D25, D27, I30, K32, S39, E41, K53, K54, I55, H68; D35, S36, K38, S39, I55, E59, D60, E62, K64, K138;	27 57	1	10

	D27, S39, E41, E44, I52, K53, K54, I55, H68, K69; K53, H68, K69, E71, S72, I73, S81, I97, D100, S101;	54 82		
KIESIDE	D35, S36, K38, S39, I55, E59, D60, E62, K64, K138;	57	1	10
IESIDEA	A21, D25, A26, D27, I30, A33, S36, S39, I55, A145; A26, D27, I30, A33, D35, S36, S39, I55, A141, A145;	30 37	1	10
ESIDEAN	A21, D25, A26, D27, N28, I30, A33, S36, S39, I55 , A145;	30	1	11
SIDEANL	A21, L22, D25, A26, D27, N28, I30, A33, L34, S36, S39, I55, L142, A145;	30	1	14
IDEANLG	G1, G86, G87, A88, A89, L90, D92, A94, E95, I97, E120, G123, A125;	94	1	13
DEANLGY	G1, G86, G87, A88, A89, L90, D92, A94, E95, Y119, E120, G123, A125;	94	1	13
EANLGYS	L22, A26, E71, S72, E75, N77, L78, G79, Y80, S81, Y82, S101; G1, G86, G87, A88, A89, L90, A94, E95, Y119, E120, G123, A125;	80 94	1	12
ANLGYSY	A21, L22, A26, N28, A33, L34, S36, S39, L142, A145, Y149;	30	1	11
NLGYSYS	L22, S72, N77, L78, G79, Y80, S81, Y82, S101;	80	1	9
LGYSYSV	V66, L67, Y82, V84, V85, G86, G87, L90, V117, Y119;	84	1	10
GYSYSVV	G50, S72, Y80, S81, Y82, V84, V85, S101; V66, Y82, V84, V85, G86, G87, V117, Y119;	82 84	1	8
YSYSVVG	G50, S72, Y80, S81, Y82, V84, V85, S101; V66, Y82, V84, V85, G86, G87, V117, Y119;	82 84	1	8
SYSVVGG	V40, V43, V66, Y82, V84, V85, G86, G87; G50, S72, Y80, S81, Y82, V84, V85, S101; V66, Y82, V84, V85, G86, G87, V117, Y119;	68 82 84	1	8
YSVVGGA	A26, V40, V43, V66, Y82, V84, V85, G86, G87;	68	1	9
SVVGGAA	G1, V84, G86, G87, A88, A89, A94, G123, A125;	94	1	9
VVGGAAL	V66, L67, V84, V85, G86, G87, A88, A89, L90, A94; V66, L67, V84, V85, G86, G87, A88, L90, A94, V117; V13, A14, A21, L22, V29, A145, A148, L150, L151, A152; L22, A26, V29, L34, G135, A137, A139, A141, L142, A145; V66, L67, V84, V85, G86, G87, A88, A89, L90, A94; A21, L22, A26, V29, A33, L34, V40, L142, A145, L150; V13, A21, V29, A33, L34, A145, A148, L150, L151, A152; G1, V84, G86, G87, A88, A89, L90, A94, G123, A125; V66, L67, V84, V85, G86, G87, A88, A89, L90, A94;	90 84 150 142 87 29 149 94 86	1	10
VGGAALP	P12, V13, A14, A21, L22, V29, A145, A148, L150, L151, A152, P154; G1, V84, G86, G87, A88, A89, L90, P91, A94, G123, A125, P127;	150 94	1	12
GGAALPD	G1, G86, G87, A88, A89, L90, P91, D92, A94, G123, A125, P127;	94	1	12

GAALPDT	G1, G86, G87, A88, A89, L90, P91, D92, T93, A94, T121, G123, A125, P127;	94	1	14
AALPDTA	A21, L22, T24, D25, A26, D27, P31, A33, L34, L142, A145, L150;	29	1	12
ALPDTAE	T4, L90, T93, A94, E95, T98, E120, T121, A125, E126, P127, L132; A88, A89, L90, P91, D92, T93, A94, E95, T121, A125, E126, P127; A21, L22, T24, D25, A26, D27, P31, A33, L34, L142, A145, L150; P12, A14, T17, A21, L22, A145, E147, A148, L150, L151, A152, P154; A88, A89, L90, P91, D92, T93, A94, E95, E120, T121, A125, P127;	119 93 29 150 94	1	12
LPDTAEK	T4, L90, T93, A94, E95, K96, T98, K118, E120, T121, A125, E126, P127, L132; K20, A21, L22, T24, D25, A26, D27, P31, K32, A33, L34, L142, A145, L150; A88, A89, L90, P91, D92, T93, A94, E95, K96, E120, T121, K122, A125, P127;	119 29 94	1	14
PDTAEKI	A88, A89, P91, D92, T93, A94, E95, K96, I97, E120, T121, K122, A125, P127;	94	1	14
DTAEKIT	T4, T93, A94, E95, K96, I97, T98, K118, E120, T121, A125, E126; A88, A89, D92, T93, A94, E95, K96, I97, E120, T121, K122, A125;	119 94	1	12
TAEKITF	F3, T4, F5, T93, A94, E95, K96, I97, T98, F99, K118, E120, T121, A125, E126;	119	1	15
AEKITFD	F3, T4, F5, T93, A94, E95, K96, I97, T98, F99, K118, E120, T121, A125, E126;	119	1	15
EKITFDS	D35, S36, F37, K38, S39, I55, T56, F57, E59, D60, E62, T63, K64, F65, K138;	57	1	15
KITFDSK	D35, S36, F37, K38, S39, I55, T56, F57, D60, T63, K64, F65, K138;	57	1	13
ITFDSKL	D35, S36, F37, K38, S39, I55, T56, F57, L58, D60, T63, K64, F65, K138; D27, I30, K32, L34, D35, S36, F37, K38, S39, I55, T56, F57, L58, L142; F3, T4, F5, D7, I97, T98, F99, D100, S101, L115, T116, K118, L132, K136;	57 37 117	1	14
TFDSKLV	F3, T4, F5, D7, V84, T98, F99, D100, S101, L115, T116, V117, K118, L132, K136;	117	1	15
FDSKLVA	A26, D27, V29, K32, A33, L34, D35, S36, F37, K38, S39, V40, F57, L58, A141, L142, A145;	37	1	17
DSKLVAG	A26, D27, V29, K32, A33, L34, D35, S36, K38, S39, V40, L58, A141, L142, A145;	37	1	15
SKLVAGP	G1, V84, G86, G87, A88, A89, L90, P91, A94, K96, K122, G123, A125, P127; A26, V29, P31, K32, A33, L34, S36, K38, S39, V40, L58, A141, L142, A145; K54, V66, L67, K69, V84, V85, G86, G87, A88, A89, L90, P91, A94, K96;	94 37 86	1	14
KLVAGPN	K20, A21, L22, A26, N28, V29, P31, K32, A33, L34, V40, L142, A145, L150; G1, V84, G86, G87, A88, A89, L90, P91, A94, K96, K122, G123, A125, P127; K54, V66, L67, K69, V84, V85, G86, G87, A88, A89, L90, P91, A94, K96;	29 94 86	1	14
LVAGPNG	P12, V13, A14, A21, L22, V29, A145, A148, L150, L151, A152, P154; N10, P12, V13, A14, P15, A16, A21, L22, G109, A148, L150, L151; A21, L22, A26, N28, V29, P31, A33, L34, V40, L142, A145, L150; G1, V84, G86, G87, A88, A89, L90, P91, A94, G123, A125, P127;	150 13 29 94	1	12
VAGPNGG	N10, P12, A14, P15, V104, A105, G106, P107, N108, G109, A112; G1, V84, G86, G87, A88, A89, P91, A94, G123, A125, P127;	106 94	1	11
AGPNGGS	N10, P12, A14, P15, A105, G106, P107, N108, G109, A112;	106	1	10

	G1, G86, G87, A88, A89, P91, A94, G123, A125, P127;	94		
GPNGGSA	N10, P12, A14, P15, A105, G106, P107, N108, G109, A112; G1, G86, G87, A88, A89, P91, A94, G123, A125, P127;	106 94	1	10
PNGGSAG	N10, P12, A14, P15, A105, G106, P107, N108, G109, A112; G1, G86, G87, A88, A89, P91, A94, G123, A125, P127;	106 94	1	10
NGGSAGK	G1, G86, G87, A88, A89, A94, K96, K122, G123, A125;	94	1	10
GGSAGKL	L22, A26, L34, S36, G135, A137, K138, A139, A141, L142, K144, A145;	142	1	12
GSAGKLT	G1, G86, G87, A88, A89, L90, T93, A94, K96, T121, K122, G123, A125;	94	1	13
SAGKLTV	G1, V2, T4, V84, L90, T93, A94, K96, T98, V117, K118, T121, A125, L132; K53, K54, T56, T63, K64, V66, L67, V84, V85, G86, G87, A88, A89, L90; G1, V84, G86, G87, A88, A89, L90, T93, A94, K96, T121, K122, G123, A125; A26, V29, K32, A33, L34, S36, K38, S39, V40, T56, L58, A141, L142, A145;	119 66 94 37	1	14
AGKLTVK	G1, V2, T4, V84, L90, T93, A94, K96, T98, V117, K118, T121, A125, L132; G1, V84, G86, G87, A88, A89, L90, T93, A94, K96, T121, K122, G123, A125; K53, K54, T56, T63, K64, V66, L67, V84, V85, G86, G87, A88, A89, L90;	119 94 66	1	14
GKLTVKY	G1, V2, T4, V84, L90, T93, K96, T98, V117, K118, Y119, T121, L132; V66, L67, K69, Y82, V84, V85, G86, G87, L90, K96, T98, V117, Y119,	119 84	1	13
KLTVKYE	V2, T4, V84, L90, T93, E95, K96, T98, V117, K118, Y119, E120, T121, E126, L132;	119	1	15
LTVKYET	V2, T4, V84, L90, T93, E95, K96, T98, V117, K118, Y119, E120, T121, E126, L132;	119	1	15
TVKYETK	V2, K69, E71, Y82, V84, V85, E95, K96, T98, T116, V117, K118, Y119, E120;	98	1	14
VKYETKG	G1, V2, T4, V84, T93, E95, K96, T98, V117, K118, Y119, E120, T121, E126; V2, K69, E71, Y82, V84, V85, E95, K96, T98, T116, V117, K118, Y119, E120;	119 98	1	14
KYETKGD	G1, G86, G87, D92, T93, E95, K96, Y119, E120, T121, K122, G123;	94	1	12
YETKGDA	G1, G86, G87, A88, A89, D92, T93, A94, E95, K96, Y119, E120, T121, K122, G123, A125;	94	1	16
ETKGDAE	G1, G86, G87, A88, A89, D92, T93, A94, E95, K96, E120, T121, K122, G123, A125;	94	1	15
TKGDAEP	G1, G86, G87, A88, A89, P91, D92, T93, A94, E95, K96, E120, T121, K122, G123, A125, P127;	94	1	17
KGDAEPN	G1, G86, G87, A88, A89, P91, D92, A94, E95, K96, E120, K122, G123, A125, P127;	94	1	15
GDAEPNQ	G1, G86, G87, A88, A89, P91, D92, A94, E95, E120, G123, A125, P127;	94	1	13
DAEPNQD	A89, A94, A125, E126, P127, N128, Q129, D130, E131; A88, A89, P91, D92, A94, E95, A125, E126, P127; A88, A89, P91, D92, A94, E95, E120, A125, P127;	127 93 94	1	9
AEPNQDE	A89, A94, A125, E126, P127, N128, Q129, D130, E131; A88, A89, P91, D92, A94, E95, A125, E126, P127; A88, A89, P91, D92, A94, E95, E120, A125, P127;	127 93 94	1	9

EPNQDEL	N10, P12, L22, D140, L142, E147, L150, L151; L22, D25, D27, N28, P31, L34, L142, L150;	147 29	1	8
PNQDELK	K20, L22, D25, D27, N28, P31, K32, L34, L142, L150;	29	1	10
NQDELKT	T4, L90, T93, E95, K96, T98, K118, E120, T121, E126, L132; T4, E6, D7, E8, T98, D100, K102, K114, L115, T116, K118;	119 116	1	11
QDELKTG	G1, T4, L90, T93, E95, K96, T98, K118, E120, T121, E126, L132; G1, G86, G87, L90, D92, T93, E95, K96, E120, T121, K122, G123;	119 94	1	12
DELKTGK	G1, T4, L90, T93, E95, K96, T98, K118, E120, T121, E126, L132; G1, G86, G87, L90, D92, T93, E95, K96, E120, T121, K122, G123;	119 94	1	12
ELKTGKA	G1, G86, G87, A88, A89, L90, T93, A94, E95, K96, E120, T121, K122, G123, A125;	94	1	15
LKTGKAK	G1, T4, L90, T93, A94, K96, T98, K118, T121, A125, L132; L22, A26, L34, G135, A137, K138, A139, A141, L142, K144, A145; L115, L132, K133, T134, G135, K136, A137, K138, A139, A141, L142;	119 142 135	1	11
KTGKAKA	G1, G86, G87, A88, A89, T93, A94, K96, T121, K122, G123, A125;	94	1	14
TGKAKAD	G1, G86, G87, A88, A89, D92, T93, A94, K96, T121, K122, G123, A125;	94	1	13
GKAKADA	G1, G86, G87, A88, A89, D92, A94, K96, K122, G123, A125;	94	1	11
KAKADAL	K20, A21, L22, D25, A26, D27, K32, A33, L34, L142, A145, L150;	29	1	12
AKADALF	L34, D35, F37, K136, A137, K138, A139, D140, A141, L142, F143, K144, A145; K32, A33, L34, D35, F37, D140, A141, L142, F143, K144, A145, A148, L150; L22, A26, L34, F37, A137, K138, A139, D140, A141, L142, F143, K144, A145; K20, A21, L22, D25, A26, D27, K32, A33, L34, F37, L142, A145, L150; A26, D27, K32, A33, L34, D35, F37, K38, F57, L58, A141, L142, A145;	141 145 142 29 37	1	13
KADALFK	L34, D35, F37, K136, A137, K138, A139, D140, A141, L142, F143, K144, A145; K32, A33, L34, D35, F37, D140, A141, L142, F143, K144, A145, A148, L150; L22, A26, L34, F37, A137, K138, A139, D140, A141, L142, F143, K144, A145; K20, A21, L22, D25, A26, D27, K32, A33, L34, F37, L142, A145, L150; A26, D27, K32, A33, L34, D35, F37, K38, F57, L58, A141, L142, A145;	141 145 142 29 37	1	13
ADALFKA	L34, D35, F37, K136, A137, K138, A139, D140, A141, L142, F143, K144, A145; K32, A33, L34, D35, F37, D140, A141, L142, F143, K144, A145, A148, L150; L22, A26, L34, F37, A137, K138, A139, D140, A141, L142, F143, K144, A145; K20, A21, L22, D25, A26, D27, K32, A33, L34, F37, L142, A145, L150; A26, D27, K32, A33, L34, D35, F37, K38, F57, L58, A141, L142, A145;	141 145 142 29 37	1	13
DALFKAI	A26, D27, I30, K32, A33, L34, D35, F37, K38, I55, F57, L58, A141, L142, A145;	37	1	15
ALFKAIE	L22, A26, I30, L34, F37, A137, K138, A139, A141, L142, F143, K144, A145, E147;	142	1	14
LFKAIEA	L22, A26, I30, L34, F37, A137, K138, A139, A141, L142, F143, K144, A145, E147;	142	1	14
FKAIEAY	F3, F5, A94, E95, K96, I97, F99, K118, Y119, E120, A125, E126; F3, Y82, A88, A89, A94, E95, K96, I97, F99, K118, Y119, E120;	119 97	1	12

KAIEAYL	A21, I30, K32, A33, L34, K144, A145, E147, A148, Y149, L150, L151, A152, Y156;	149	1	14
AIEAYLL	A21, I30, A33, L34, A145, E147, A148, Y149, L150, L151, A152, Y156;	149	1	12
IEAYLLA	A21, I30, A33, L34, A145, E147, A148, Y149, L150, L151, A152, Y156;	149	1	12
EAYLLAH	A21, A33, L34, A145, E147, A148, Y149, L150, L151, A152, H153, Y156; A14, A21, L22, A145, E147, A148, Y149, L150, L151, A152, H153, Y156;	149 150	1	12
AYLLAHP	P12, A14, A21, L22, A145, A148, Y149, L150, L151, A152, H153, P154, Y156;	150	1	13
YLLAHPD	P12, A14, A21, L22, A145, A148, Y149, L150, L151, A152, H153, P154, Y156;	150	1	13
LLAHPDY	P12, A14, A21, L22, A145, A148, Y149, L150, L151, A152, H153, P154, Y156;	150	1	13
LAHPDYN	P12, A14, A21, L22, A145, A148, Y149, L150, L151, A152, H153, P154, Y156; A21, L22, D25, A26, D27, N28, P31, A33, L34, L142, A145, Y149, L150;	150 29	1	13

152 peptides with an off-set of one amino acid were generated from Gly m 4 amino acid sequence and mapped onto Gly m 4 surface using EpiSearch resulting in 270 potential IgE-binding mimotopes. Resulting from EpiSearch analysis, center residue, score and total length of each identified mimotope are listed.

Table A 3: Summary of results obtained with dot blot with eleven ΔNCS$_{N42/P49}$ variants (ΔNCS$_{N42/P49}$_1-_11) and patients' sera.

	Generated ΔNCS$_{N42/P49}$ variants (ΔNCS$_{N42/P49}$_x)										
Patient No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1				X		X	X	X	X	X	
2				X		X	X		X	X	
3							X		X	X	
4				X	X	X	X		X	X	
5							X				
7							X		X	X	
8				X			X		X		
9							X		X	X	
10							X	X			
11				X		X	X	X		X	
13	X				X	X	X	X			
22							X				
28										X	
30							X	X	X	X	
33							X	X	X	X	X
34			X				X				
35				X		X	X	X	X	X	X
36				X			X		X	X	
37				X			X		X	X	
38				X		X	X		X	X	
40				X			X		X	X	
41				X		X	X		X	X	
16				X		X	X		X	X	

20				X			X			X	
24				X			X			X	
25				X			X			X	
26				X			X			X	
27							X			X	
43				X			X			X	
44				X			X			X	
45										X	
46							X				
47											
48							X			X	
49							X				
50							X				

Patients sensitized and clinically soy allergic (group 1) as well as patients sensitized but without soybean allergy (group 2) were analyzed in dot blot together with generated $\Delta NCS_{N42/P49}$ variants. X resembles a positive IgE signal as described in 3.5.2. $\Delta NCS_{N42/P49}$_x resembles:

$\Delta NCS_{N42/P49}$_1=$\Delta 51NCS_{F3/F5/E6N42/P49/D124/E126}$,
$\Delta NCS_{N42/P49}$_2=$\Delta 51NCS_{D25/D27/N28/V29/K32/N42/P49/A152/D155/Y156}$,
$\Delta NCS_{N42/P49}$_3=$\Delta 51NCS_{K38/S39/N42/P49/E59/D60/K64}$,
$\Delta NCS_{N42/P49}$_4=$\Delta 51NCS_{N42/E44/N46/P49/K54}$,
$\Delta NCS_{N42/P49}$_5=$\Delta 51NCS_{N10/N42/P49/D130/E131/K133/K136/K144/E147/A148}$,
$\Delta NCS_{N42/P49}$_6=$\Delta 29NCS_{K32/N42/E44/N46/P49/K54/D60/E120/E126}$,
$\Delta NCS_{N42/P49}$_7=$\Delta 51NCS_{F3/K38/N42/P49/D60/N128}$,
$\Delta NCS_{N42/P49}$_8=$\Delta 51NCS_{T24/S36/N42/P49/S81/D124}$,
$\Delta NCS_{N42/P49}$_9=$\Delta 51NCS_{T4/T17/N42/N46/P49/S83}$,
$\Delta NCS_{N42/P49}$_10=$\Delta 51NCS_{F5/T17/T24/S36/N42/P49/K64/S81/K118/D124/K133}$,
$\Delta NCS_{N42/P49}$_11=$\Delta 51NCS_{T4/T17/K32/N42/N46/P49/E75/S83\ /D100/D124/E126/K144}$.

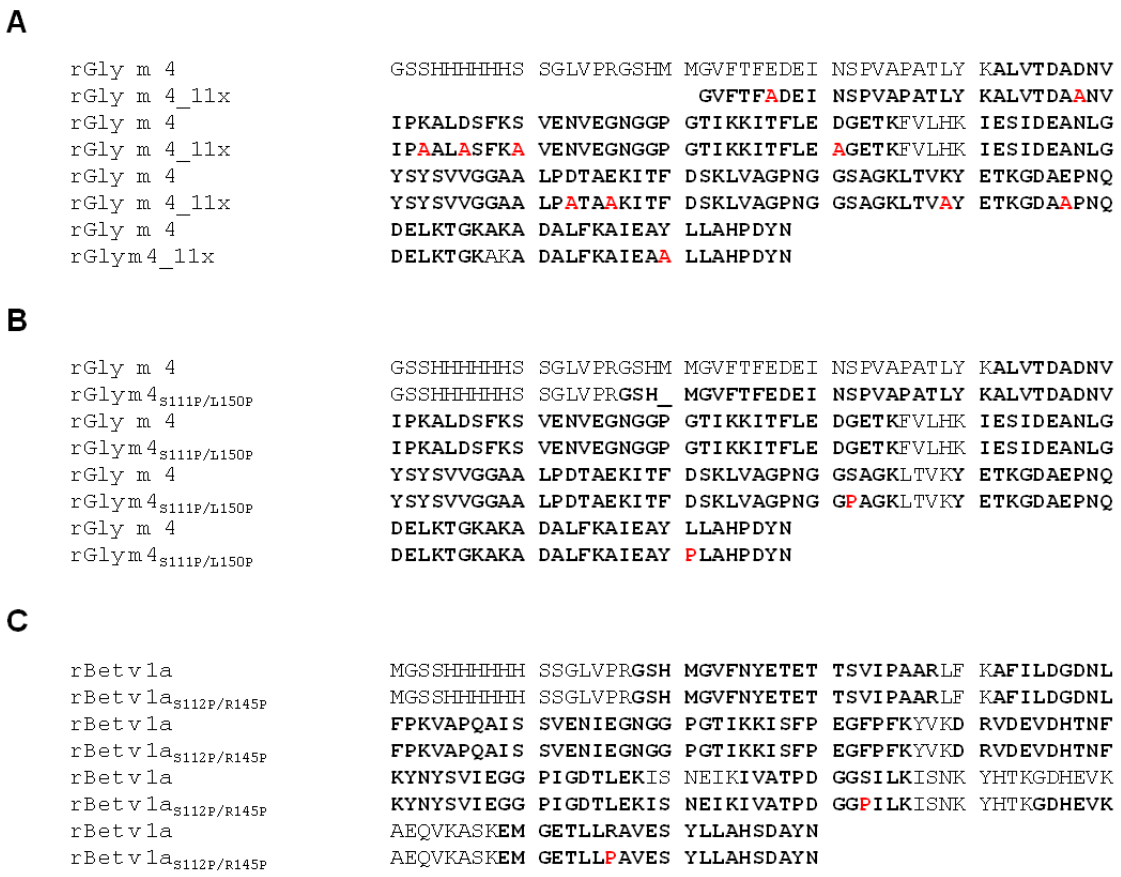


Figure A 1: Sequence alignments of proteins analyzed with LC-MS.

(A) Gly m 4 wild type (UniProt accession number P26987) and Gly m 4_11x; (B) Gly m 4 wild type (UniProt accession number P26987) and Gly m $4_{S111p/L150P}$; (C) Bet v 1a (UniProt accession number P15494) and Bet v $1a_{S111P/R145P}$; Peptides that were detected by MS are printed in bold letter. Amino acids that were exchanged are printed in red color.

Table A 4: Summary of results from actual and theoretical values of quantitative IgE immunoassays.

Immunoblot

rBet v 1a ($_{S112P/R145P}$) [%]	mean estimate (actual) in [AU]	95% CI	p-value	mean estimate (theoretical) in [AU]	actual/theoretical comparison
0	$1.17*10^{7}$	$6.56*10^{6}$-$2.1*10^{7}$	--	$1.17*10^{7}$	1
20	$8.86*10^{6}$	$4.95*10^{6}$-$1.58*10^{7}$	0.0088	$9.39*10^{6}$	0.94
40	$6.88*10^{6}$	$3.85*10^{6}$-$1.23*10^{7}$	0.0001	$7.04*10^{6}$	0.98
60	$5.7*10^{6}$	$3.19*10^{6}$-$1.02*10^{7}$	<0.0001	$4.69*10^{6}$	1.21
80	$3.47*10^{6}$	$1.94*10^{6}$-$6.21*10^{6}$	<0.0001	$2.35*10^{6}$	1.48
90	$2.04*10^{6}$	$1.14*10^{6}$-$3.66*10^{6}$	<0.0001	$1.17*10^{6}$	1.74
99*	$2.55*10^{4}$	$1.43*10^{4}$-$4.57*10^{4}$	<0.0001	$1.17*10^{5}$	0.22
99.9	$4.72*10^{4}$	$2.65*10^{4}$-$8.46*10^{4}$	<0.0001	$1.17*10^{4}$	4.03
99.99	$1.54*10^{4}$	$8.61*10^{3}$-$2.75*10^{4}$	<0.0001	$1.17*10^{3}$	13.16
100	$1.56*10^{4}$	$8.75*10^{3}$-$2.8*10^{4}$	<0.0001	--	--

Inhibition ELISA

rBet v 1a ($_{S112P/R145P}$) [%]	actual EC_{50} [ng/well]	95% CI	p-value	theoretical EC_{50} [ng/well]	actual/theoretical comparison
0	2.41	1.97-2.96	--	2.41	1
20	3.37	2.75-4.15	0.588	3.01	1.09
40	3.75	3.05-4.62	0.412	4.02	0.93
60	5.99	4.83-7.44	0.182	6.03	0.99
80	10.46	8.32-13.2	0.039	12.05	0.87
90	17.31	13.55-22.2	0.093	24.1	0.72
99	144.45	109.57-190.73	0.724	241	0.61

99.9	691.05	503.35-954.87	0.002	2410	0.29
99.99	no slope	--	--	24100	--
100	no slope	--	--	--	--

Mediator release (RBL)

rBet v 1a ($_{S112P/R145P}$) [%]	**actual EC_{50} [ng/well]**	**95% CI**	**p-value**	**theoretical EC_{50} [ng/well]**	**actual/theoretical comparison**
0	0.68	0.4-1.17	--	0.68	1
20	0.73	0.43-1.25	0.995	0.85	0.86
40	0.86	0.51-1.51	0.984	1.13	0.76
60	0.81	0.48-1.4	0.951	1.7	0.48
80	0.9	0.53-1.58	0.902	3.4	0.26
90	2.05	1.08-3.87	0.936	6.8	0.3
99	89.76	44.4-174.8	0.682	68	1.32
99.9	no slope	--	--	680	--
99.99	no slope	--	--	6800	--
100	no slope	--	--	--	--

Shown are experimental values, 95% confidence intervals, p-values, theoretical EC_{50} values and actual/theoretical comparisons for immunoblot, inhibition ELISA and mediator release (RBL).

8.2 List of figures

8.3 List of tables

9 PUBLIKATIONSLISTE

Peer reviewed

Felix Husslik, Jasmin Nürnberg, Christian Seutter von Loetzen, Timo Mews, Barbara Ballmer-Weber, Jörg Kleine-Tebbe, Regina Treudler, Jan-Christoph Simon, Stefanie Randow, Elke Völker, Andreas Reuter, Paul Rösch, Stefan Vieths, Thomas Holzhauser, Dirk Schiller.
The conformational IgE epitope profile of soya bean allergen Gly m 4.
Clin Exp Allergy. 2016 Nov;46(11):1484-1497. doi: 10.1111/cea.12796

Felix Husslik, Kay-Martin Hanschmann, Ariane Krämer, Christian Seutter von Loetzen, Kristian Schweimer, Iris Bellinghausen, Regina Treudler, Jan C. Simon, Lothar Vogel, Elke Völker, Stefanie Randow, Andreas Reuter, Paul Rösch, Stefan Vieths, Thomas Holzhauser, Dirk Schiller.
Folded or Not? Tracking Bet v 1 Conformation in Recombinant Allergen Preparations.
PLoS ONE. 2015; 10(7): e0132956. doi: 10.1371/journal.pone.0132956

Hanna Berkner, Christian Seutter von Loetzen, Maximilian Hartl, Stefanie Randow, Michaela Gubesch, Lothar Vogel, **Felix Husslik**, Andreas Reuter, Jonas Lidholm, Barbara Ballmer-Weber, Stefan Vieths, Paul Rösch, Dirk Schiller.
Enlarging the Toolbox for Allergen Epitope Definition with an Allergen-Type Model Protein.
PLoS ONE. 2014; 9(10): e 111691. doi: 10.1371/journal.pone0111691

Kongressbeiträge

Vorträge

26. Mainzer Allergie Workshop, DGAKI, Universitätsmedizin Mainz, 6.-7. März 2014

PEI-Retreat 2014, Paul-Ehrlich-Institut, Heidelberg, 15.-17. Januar 2014

24. Mainzer Allergie Workshop, DGAKI, Universitätsmedizin Mainz, 22.-23. März 2012

Posterpräsentationen

International Symposium on Molecular Allergology, EAACI, Wien, Österreich, 5.-7. Dezember 2013

PEI-Retreat 2013, Paul-Ehrlich-Institut, Altenhau, 18.-20. Januar 2013

PEI-Retreat 2012, Paul-Ehrlich-Institut, Altenhau, 13.-15. Januar 2012

10 LEBENSLAUF

Persönliche Informationen

Name, Vorname:	Husslik, Felix Michael
Geburtsdatum:	01. April 1987
Geburtsort:	Miltenberg
Staatsangehörigkeit:	Deutsch

Berufliche Tätigkeit

01/2016 - heute	Junior GMP Consultant, NNE Pharmaplan, Bad Homburg

Akademische Ausbildung

11/2011 - 11/2014	Promotionsarbeit im Fachgebiet Rekombinante Allergentherapeutika, Paul-Ehrlich-Institut, Langen
10/2009 - 09/2011	Masterstudiengang Biochemie und molekulare Biologie, Universität Bayreuth Abschluss: Master of Science, Biochemie und molekulare Biologie
10/2006 - 09/2009	Bachelorstudiengang Biochemie, Universität Bayreuth Abschluss: Bachelor of Science, Biochemie

Schulausbildung

09/1997 – 07/2006	Johannes-Butzbach Gymnasium, Miltenberg Abschluss: Allgemeine Hochschulreife
09/1993 – 07/1997	Grundschule Großheubach

Berufliche Weiterbildung, Praktika

06/2015 - 07/2015	zertifiziertes Qualifizierungsseminar „GMP für Akademiker“, i.b.s. human resources, Jena

Sprachkenntnisse

Deutsch:	Muttersprache
Englisch:	fließend
Französisch:	Grundkenntnisse

_____

11 ERKLÄRUNGEN

Felix Husslik
Mendelstraße 11
63814 Mainaschaff

_____

Erklärung

Ich erkläre hiermit, dass ich meine Dissertation selbstständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

_____

Felix Husslik
Mendelstraße 11
63814 Mainaschaff

_____

Erklärung

Ich erkläre hiermit, noch keinen Promotionsversuch unternommen zu haben.

_____

12 DANKSAGUNG

An erster Stelle möchte ich mich bei Prof. Dr. Stefan Vieths und Dr. Thomas Holzhauser für die Bereitstellung dieses interessanten Themas bedanken sowie für die Möglichkeit in der Abteilung Allergologie des Paul-Ehrlich-Instituts an diesem Projekt arbeiten zu dürfen.

Weiterhin ein großer Dank an Prof. Dr. Harald Kolmar für die Übernahme der externen Betreuung von Seiten der TU Darmstadt.

Besonders danken möchte ich Dr. Dirk Schiller für die durchgehend professionelle und sehr lehrreiche Begleitung der kompletten Arbeit. In dieser Zeit standst du mir immer für Fragen und Diskussionen zur Verfügung und deine Unterstützung trug enorm zum Gelingen dieser Forschungsarbeit bei.

Danken möchte ich auch allen klinischen Partnern für die Durchführung der klinischen Untersuchungen und die Bereitstellung der Patientenseren. Besonders sind dies PD Dr. Jörg Kleine-Tebbe vom Allergie- und Asthmazentrum Berlin, Prof. Dr. Barbara Ballmer-Weber vom Universitätskrankenhaus Zürich, PD Dr. Regina Treudler von der Universität Leipzig und Dr. Iris Bellinghausen von der Universität Mainz.

Für die Durchführung und Auswertung der MS-Analysen sowie Einblicke in die Massenspektrometrie möchte ich mich bei Luisa Schwaben und Dr. Andreas Reuter bedanken.

Bei Dr. Christian Seutter von Loetzen, Prof. Dr. Paul Rösch und Dr. Kristian Schweimer bedanke ich mich für die unkomplizierte Zusammenarbeit, hilfreichen Diskussionen und vor allem für die Überprüfung der Proteine mittels NMR-Spektroskopie.

Kay-Martin Hanschmann danke ich für die Unterstützung bei allen statistischen Auswertungen.

Dank sagen möchte ich auch Dr. Lothar Vogel für die Anleitung und Durchführung der RBL-Assays.

Ein großer Dank geht an Michaela Gubesch für die Durchführung der ImmunoCAP-Messungen.

Dank gilt allen Angehörigen der Abteilung 5 des Paul-Ehrlich-Instituts für den spannenden Laboralltag und die angenehme Arbeitsatmosphäre. Besonders hervorheben möchte ich Jasmin Nürnberg, Elke Völker, Dr. Tom Schulenborg, Stefanie Randow, Dr. Jelena Spiric, Ariane Krämer, Mike Hellmuth, Anna Koutsouridu, Andrea Wangorsch, Kirsten Kuttich, Marie Messmer und Dr. Stefan Schülke. Vielen Dank für die zahlreichen lustigen und heiteren Momente am PEI und bei diversen Freizeitaktivitäten. Ihr werdet mir ganz besonders fehlen und ich denke gerne an unsere gemeinsame Zeit am PEI zurück.

Zu guter Letzt ein riesengroßes Dankeschön an meine Schwester Anna, meine Eltern und besonders an Alisha dafür, dass ihr immer für mich da seid, mich unterstützt und immer an mich geglaubt habt. Ohne euch wäre das alles nicht möglich gewesen.