Judith Wortmann, BSc

Structural and biophysical characterization of the four chimeric allergen constructs Api-Bet 1-4

MASTER'S THESIS

to achieve the university degree of

Master of Science

Master's degree programme: Biochemistry and Molecular Biomedical Sciences

submitted to

Graz University of Technology

Supervisor

Ao.Univ.-Prof. Mag. Dr.rer.nat. Walter Keller

Institute of molecular biosciences

Graz, December 2015

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Date

Signature

Table of content

I Danksagung

An erster Stelle möchte ich mich bei meinem Betreuer Professor Walter Keller bedanken, dass ich meine Masterarbeit bei ihm absolvieren durfte. Ich bin dankbar dafür, dass ich im Rahmen dieses Großprojektes so viele interessante Erfahrungen in der wissenschaftlichen Welt machen konnte, mein eigenes Projekt selbstständig bearbeiten konnte und bei Problemen immer eine offene Tür vorgefunden habe!

Danke auch an Barbara Gepp von der Medizinischen Universität Wien, die mir die Konstrukte und Plasmide zur Verfügung gestellt hat

Mein besonderer Dank gilt außerdem meinem Kollegen Gerhard Hofer, der mir alles nötige im Labor beigebracht hat, stets geduldig und hilfsbereit war und mich mit seinem riesigen Wissen immer wieder beeindruckt hat. Ich möchte ihm und auch Markus Eder für die tolle Zusammenarbeit und die interessanten und lustigen Diskussionen danken.

Danke an Christian Fercher, der die SAXS-Messung durchgeführt hat und mir beim Datenprozessieren und auch darüber hinaus häufig geholfen hat.

Natürlich danke ich der gesamten „Strubi“-Gruppe für die tolle Arbeitsatmosphäre und die Hilfsbereitschaft

Ich danke meinen lieben Wohnungskollegen Clemens und Gregor Schöffl für eine schöne gemeinsame Zeit in Graz abseits vom Studium!

Ein ganz besonderer Dank gilt meinem Freund Martin Schwarzbach, der besonders während der Entstehung dieser Arbeit stets an meiner Seite war und auch die schwierigen Phasen mit mir durchgestanden hat und sich meine Ausführungen zu meiner Arbeit immer geduldig angehört hat. Außerdem ein Danke für die Hilfe beim Datenfitting.

Ganz zum Schluss möchte ich meinen lieben Eltern Mechthild und Gregor von ganzem Herzen dafür danken, dass sie mich immer und bedingungslos in allen Lebenslagen unterstützt haben und stets geduldig waren. Ohne ihre frühe Förderung und ihren Einsatz hätte ich diese Laufbahn wohl nicht einschlagen können. Danke!

II List of Abbreviations

AB	Api-Bet chimeric constructs
Ab	Antibody
APC	Antigen presenting cell
CV	column volume
CD	Circular Dichroism
DC	Dendritic cell
ddH_2O	double distilled water
E. coli	Escherichia coli
Fab	antigen binding domains in antibodies
Fc	constant domains in antibodies
FcεRI/II	high/low affinity receptor for IgE
FEM	feature enhanced map
FT	flowthrough
His-tag	Histidine-tag
IB	inclusion body
IgE	Immunoglobulin E
IL	Interleukin
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl-β-thiogalactopyranosid
kDa	kilo Dalton
LB	Lysogency Broth media
M	unstained protein marker

MHC	Major histocompatibility complex
MR	molecular replacement
MW	molecular weight
NMR	Nuclear magnetic resonance
OD_{600}	optical density at 600 nm
ONC	over night culture
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PGLB	protein gel loading buffer
Rpm	revolutions per minute
RT	room temperature
SAXS	small angle X-ray scattering
SCIT	allergen-specific subcutaneous immunotherapy
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
SIT	allergen-specific immunotherapy
SLIT	allergen-specific sublingual immunotherapy
TCR	T-cell receptor
T_H	T helper cell
THX	thioredoxin
Tris	2-Amino-2-hydroxymethyl.propane-1,3-diol
UV	ultraviolet (light)

III Abstract

Allergic disorders have been on the rise over the last decades which makes the research on this field even more important, especially on the mechanisms like the IgE-epitope investigation in type I-hypersensitivities.

A new technique to map conformational IgE-epitopes is the chimera design. For the mapping of these epitopes of the major birch pollen allergen Bet v 1, four chimeric constructs were designed based on the scaffold of Api g 1, the Bet v 1 homologue in celery, with 10-13 mutations to Bet v 1 specific surface residues (Api-Bet 1-4).

The aim of the study was the biophysical and structural characterization of these four chimeric constructs to investigate the impact of the mutations on the overall fold. First all chimeras were expressed in Escherichia coli in different setups and strategies for the purification of the four proteins were established. The structural integrity and the temperature stability were tested with CD-spectroscopy. Crystallization trails of each chimera were performed under varying conditions.

So far, one crystal structure of Api-Bet 1 could be solved that shows an unexpected, domain-swapped and dimeric arrangement. This structure strongly deviates from the expected monomeric Bet v 1-like fold of the parent protein structures.

To evaluate this effect found in the crystal structure of Api-Bet 1, further research has to be done, also regarding the other three chimeras.

IV Zusammenfassung

Allergien sind in den letzten Jahrzehnten zu einem immer größer werdenden Problem geworden, was die Forschung in diesem Bereich immer wichtiger werden lässt. Ein wichtiger Aspekt ist dabei die Untersuchung der IgE-Allergen Bindung.

Eine neue Technik zur Bestimmung von konformationellen IgE-Epitopen ist das Design von Chimärkonstrukten. Zur Bestimmung dieser Epitope des Hauptallergens in Birkenpollen, Bet v1, wurden vier chimäre Konstrukte entwickelt, basierend auf der Faltung von Api g 1, dem Homolog zu Bet v 1 aus Sellerie mit 10-13 Mutationen zu dem entsprecheneden Oberflächenaminosäuren von Bet v 1 (Api-Bet 1-4)

Das Ziel dieser Arbeit war die biophysikalische und strukturelle Charakterisierung dieser vier chimären Konstrukte, um den Einfluss der Mutationen auf den Gesamtfaltungszustand zu untersuchen. Zunächst wurden alle vier Chimären als verschiedene Konstrukte in E. coli exprimiert und anschließend Strategien zur Proteinreinigung etabliert. Die strukturelle Integrität und die Temperaturstabilität der vier Proteine wurde mittels CD-Spektroskopie untersucht. Kristallisationsansätze zur röntgenkristallographischen Strukturaufklärung wurden in verschieden Bedingungen angesetzt.

Bisher konnte eine Kristallstruktur von Api-Bet 1 gelöst werden. Diese Struktur zeigt einen unerwarteten Domänenaustausch, der zu einer Dimerisierung geführt hat. Die gelöste Struktur weicht stark von der erwarteten monomeren Bet v 1-like Faltung der Elternproteine ab.

Zur Beurteilung dieses Effektes in der Kristallstruktur werden weitere Untersuchungen benötigt, auch bezüglich der anderen drei Chimären.

1 Introduction

This present master's thesis work was performed during a large cooperation project on allergy research with the focus on the prevention and therapy of allergy. Whereas the work in the group of structural biology at the University of Graz is focused on the expression, purification and characterization of allergens and allergy-related proteins and constructs as well as on the development of new computational approaches, the first part of the present work should give the theoretical background on allergy and its mechanisms in which the practical work is embedded.

1.1 Allergy

The field of allergy-related diseases includes a variety of disorders triggered by a hypersensitivity of the immune system. Nevertheless, not every hypersensitive immune reaction is caused by an allergic reaction. The Asthma and Allergy Foundation of America (AAFA) defines allergy as follows:

> "Allergy is characterized by an overreaction of the human immune system to a foreign protein substance ('allergen') that is eaten, breathed into the lungs, injected or touched. This immune overreaction can result in symptoms such as coughing, sneezing, itchy eyes, runny nose and scratchy throat. In severe cases it can also result in rashes, hives, lower blood pressure, difficulty breathing, asthma attacks, and even death." (www.aafa.org, *Allergy Facts and Figures*, September 2015)

Therefore, allergy is a hypersensitivity reaction towards a practically harmless substance. In the last decades, allergies have become an increasing problem in industrialized countries with 25-30% of the people affected [1]. A study of 2007 has shown that 64% of the children living in a polish city are sensitized to at least one allergen (skin prick test) [2].

The causation of allergy is still not fully clarified, whereas beside other factors, there is a correlation between the cumulative occurrence of hypersensitivities and modern lifestyle with increased hygiene in 'Western' countries [3]. It seems that allergic disorders will rise in the future and this problem will become even more relevant which makes this field of research evermore necessary.

The molecular mechanisms of hypersensitivities, the resulting symptoms as well as the methods of diagnosis and the current treatment are explained in the following chapters.

1.1.1 Types of hypersensitivities

Allergies are a set of hypersensitivities that deviate in their clinical manifestation and in their immunological reactions and mediators. While there are other forms of hypersensitivity, like autoimmunity, the allergic hypersensitivities can be classified into four groups. This classification was introduced by Gell and Coombs based on the immunological responses and the mediators involved and is still in use [4].

Type I hypersensitivity also called immediate type is an IgE-mediated reaction that leads to mast cell activation when specific IgE-Abs produced by plasma cells after a previous sensitization bind to soluble antigens. The detailed mechanisms of this type are explained in detail in the following chapter.

Type II is an IgG or IgM-mediated reaction against cell-surface or matrix-associated allergens and leads to opsonisation. In this form of hypersensitivity the complement system is activated and FcR+ cells like NK cells and Phagocytes are recruited.

Type III is also mediated by IgG-Abs and leads to complement activation via complex formation of Antibodies and soluble Antigens in several tissues that causes inflammation. This process is enhanced by the infiltration of neutrophiles.

Type IV hypersensitivity is the only type that is T-cell mediated and is also called delayed-type. In this type, T helper cells secrete pro-inflammatory cytokines after the contact with soluble antigens that lead to an inflammation process. Another mechanism is CTL-mediated cytotoxicity against cell-associated antigens. An example for this reaction is the contact dermatitis.

1.1.2 Type-I hypersensitivity

The clinically most important type of allergy is the IgE-mediated type-I. This allergic disorder can be divided into three phases: the sensitization phase, the immediate reaction and a late phase reaction. The molecular mechanisms are explained in detail in the following chapters. The major characteristic of this type of hypersensitivity is the production of specific IgE-antibodies by allergic patients towards an allergic substance after a sensitization to it. Instead of tagging a foreign substance and removing it, the binding of IgE-antibodies leads to a signal transduction cascade in mast cells and basophile granulocytes and therefore in the release of mediator substance that cause the typical symptoms [5].

1.1.2.1 Sensitization and memory phase

The first step in developing a pathophysiological allergic mechanism is the sensitization to an allergic substance. When an allergen enters the body, it gets incorporated by dendritic cells. These cells are antigen-presenting cells (APCs) that process antigenic proteins and peptides and present small peptide fragments via their MHC-II-complex on the surface to naive T-cells. The interaction of the MHC-complex and the T-cell receptor (TCR) by the allergic peptide with a T-cell epitope leads to the differentiation of the naive T-cell to a CD4+-T_H2-cell. On the other hand, B-cells can also get in contact with the allergen. If additionally an activated, allergen-specific T_H2-cell interacts with this B-cell via their receptors and by the stimulation of secreted cytokines (IL-4, IL-13), the antibody-class switching is induced and the B-cell differentiates to an IgE-secreting plasma cell (Fig. 1).

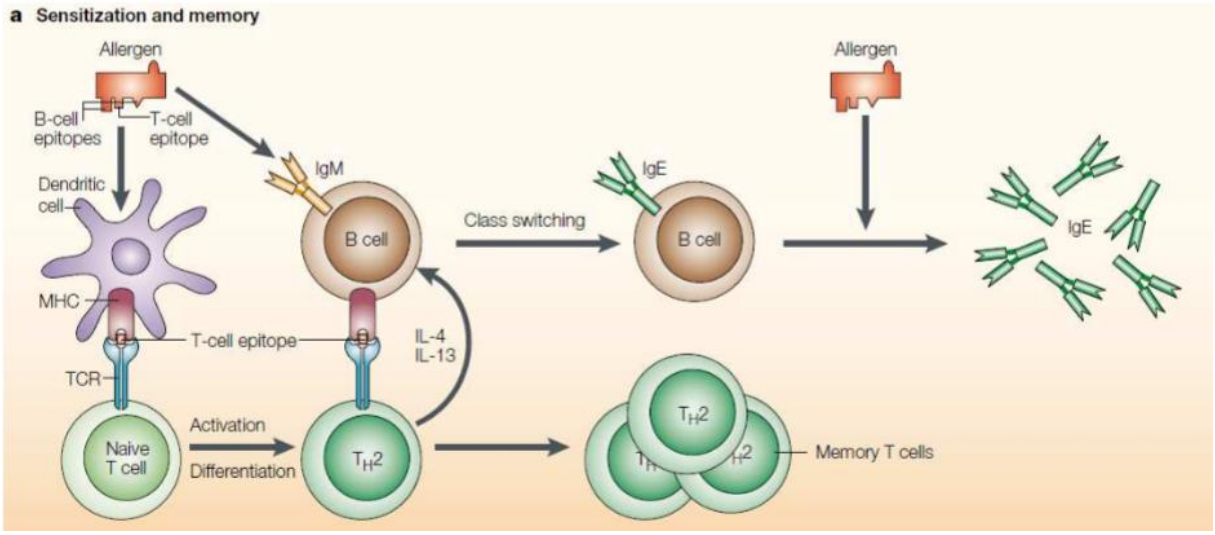


Figure 1: *The mechanism of the sensitization and memory in the type-I hypersensitivity* **[5]**

So an allergenic substance has to have a T-cell epitope that gets recognized after the processing of dendritic cells, as well as a B-cell epitope. The B-cell-T-cell interaction is crucial for the production of large amounts of IgE. This antibody isotype is responsible for the immediate reaction to a following exposure to the same allergen. The activated T_H2-cells now have memory function and can induce the differentiation of more B-cells [6, 7].

1.1.2.2 Immediate reaction phase

When a sensitized person gets in contact again with the allergenic substance, a reaction cascade starts immediately after the allergen crosses the mucosal epithelium resulting in severe tissue and organ specific symptoms. The allergen-specific IgE-antibodies that were produced by plasma cells after the sensitization phase are now bound to the highly affine FcεRI-receptors on the surface of mast cells and basophiles. Binding of the allergen to one IgE-molecule does not result in any signal transduction. As soon as an allergen gets bound by two IgE-molecules simultaneously, they get close to each other as well as both bound FcεRI which leads to cross-linking of the receptors. This cross-linking triggers a signal transduction that induces the release of the stored granula in the interior of the mast cell, called degranulation (Fig. 2). These granula are containing histamines and other mediators that lead to a local inflammatory reaction and therefore to the symptomatic manifestation of allergy. To sum up, a primary sensitization to an allergen is crucial for an immediate reaction at a following exposure and an allergen has to be recognized by two IgE-molecules at the same time to get an allergic immune reaction started [6, 8].

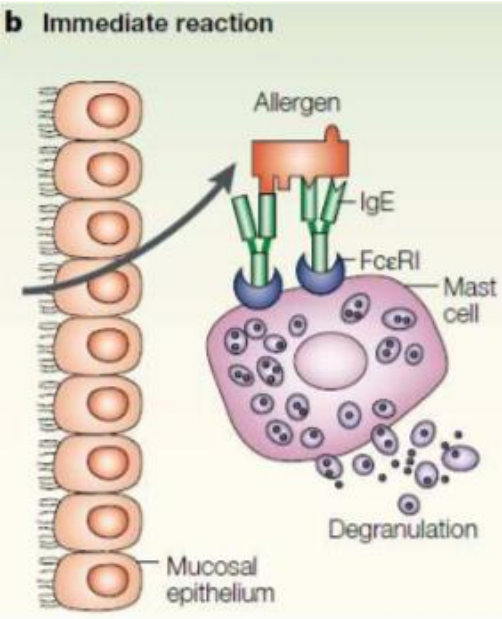


Figure 2: *The immediate allergic reaction to a renewed allergen exposure after sensitization* **[5]**

1.1.2.3 Late phase reaction

The late phase of type-I hypersensitivity is characterized by chronic inflammation and tissue damage. This reaction is caused by T_H2-cells that are recruited to the site of allergen-exposure by pro-inflammatory cytokines and mediators. Now APCs, like B-cells, macrophages and dendritic cells (IgE-facilitated and non-IgE-facilitated cells) can activate these T_H2-cells via allergen presentation resulting in the proliferation of this type of lymphocytes. Beside the secretion of IL-4 and IL-13 (as mentioned before), T_H2-cells release the pro-inflammatory homodimeric cytokine IL-5 that recruits further cells of the immune system and is especially involved in the maturation, differentiation, as well as in the effector functions of eosinophile granulocytes in the blood and local tissues (Fig. 3). Eosinophiles are the key players in a chronic late allergic reaction. The degranulation with the release of the major basic protein, peroxidases and other mediators leads to an enhanced inflammation and therefore, tissue damage [6, 9].

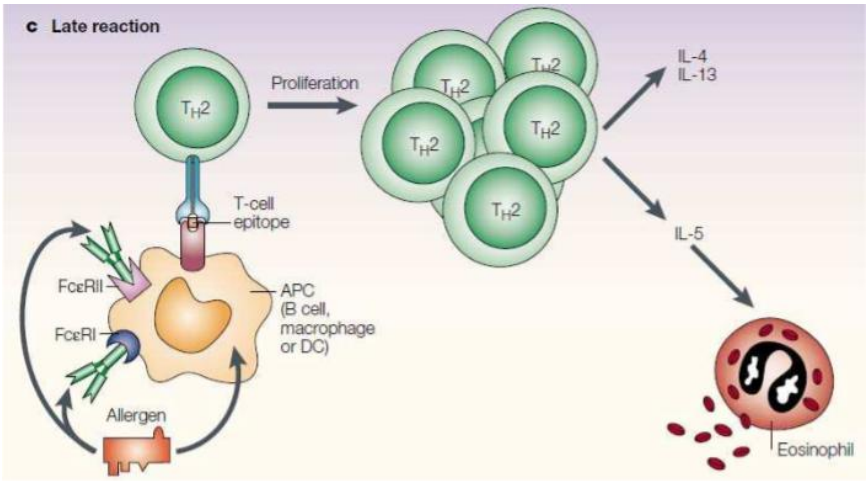


Figure 3: *The late phase reaction in type-I hypersensitivity leading to chronic inflammation* **[5]**

1.1.2.4 Symptoms

Because the effect of an allergic reaction is a local inflammation process, the resulting symptoms of allergy depend on the site of the allergen exposure. The mast cells are especially located in the mucosa, the site where foreign substances can enter the body.

Aeroallergens that get inhaled, like pollen and dust mite feces, can lead to allergic rhinitis by causing an irritation of the nasal mucosa and edema as well as to allergic asthma via bronchial constriction and increased mucus production. Food allergens that get ingested like milk, wheat or fruits can cause an irritation of the oral mucosa and also gastro-intestinal symptoms like vomiting and diarrhea. Insect bites and allergy testing (skin prick test) lead to local edema, swelling and vasodilatation of the skin.

The most fatal reaction to an allergen exposure is the systemic anaphylaxis. The clinical manifestation of such a scenario is a vasodilatation and edema, as well as a swelling of the tracheal mucosa with an occlusion. Furthermore it can cause a circulatory collapse and even death. Anaphylactic reactions especially result from intravenous exposures to drugs and venoms [10].

1.1.3 Immunoglobulin E in allergy

As explained in the chapter about the molecular mechanism of type-I hypersensitivity, the immunoglobulin E is the key mediator in the allergic immediate reaction. Like the other four classes of immunoglobulins, IgE share a similar structure, consisting of two heavy and two light chains. Beside the two antigen-binding Fab domains, the Fc region of the antibody is responsible for receptor binding. Compared to other antibodies, the concentration level of IgE in the circulation is very low (0.05 % of the IgG concentration). The reason for this phenomenon is a high-affinity receptor on the surface of mast cells that binds free IgE even when no antigen is bound to the antibody [10].

There are two Signal transduction receptors for IgE, the high affinity FcεRI and the low affinity receptor FcεRII (CD23). Binding of these receptors to the Fc region of IgE does not induce a cellular effect. But the binding of two IgE molecules to one antigen can reveal an aggregation of the bound receptors on the cell surface. This can induce severe signal transduction pathways with different effector responses. Additionally, the expression of the FcεRI receptor is increased. Though, the binding of IgE to its receptor and the caused signal transduction is critical for the manifestation of an allergic disorder [11, 12].

The physiological role of the IgE-mediated immune reaction is the defense against parasitic infections. Therefore, the cause of allergies can be seen as the effect of less parasitic infections resulting in an overreaction towards normally harmless antigens [10].

Additionally, the epitope recognition of IgE deviates from that of IgG antibodies which is further explained in the following chapter.

1.1.4 Cross-reactivity

As already mentioned, IgE-antibodies are the important players in type-I allergy. These molecules are responsible for the immediate allergic reaction by binding their target allergen after the sensitization to it. In comparison to IgG-antibodies that bind the specific antigen by the recognition of a sequential epitope, IgE-molecules are recognizing conformational as well as linear epitopes of proteins and

peptides. If allergens are structurally highly conserved and sequentially similar, they can be recognized by the same IgE-molecule even when the sensitization process took place only to the original protein. So an immediate allergic reaction can also result in the binding of two IgE-antibodies to a protein homologous to the sensitizing allergen. This phenomenon is called cross-reactivity or cross-reactive allergy. One example of this mechanism is the birch pollen allergy. A sensitization to the major birch pollen allergen Bet v 1 that has many homologues in plant foods can lead to an allergic reaction towards certain fruits and vegetables in affected people [13].

1.1.5 Diagnosis of allergy

The most common way to test for an allergic disorder is the skin prick test with different allergen extracts. This testing is a challenge test where the immediate symptomatic reaction against the allergens to which a patient is sensitized is detected. Another important method is the detection of allergen-specific IgE-Abs from blood sera by testing of the IgE-binding to their specific immobilized allergen. This technique is also available on a chip-based microarray setup (ISAC) [14]. Also important in the screening for type-I allergies is the basophile activation assay (BAT).

1.1.6 Therapy of allergy

Till today, there is no cure of Type-I allergy available. Because allergy is a complex disease with many mediators involved and varying symptoms, there are different approaches of treatment. The first step of treatment is the avoidance of the allergen exposure. However, in many cases this is not possible. The symptomatic treatment can be performed with mast cell stabilizers and H1-antihistamines as well as anticholinergics in an oral application. Additionally, corticosteroids in a local application can reduce the inflammation process by the immune repressive affect.

A possibility for a targeted therapy is the monoclonal antibody against human IgE at the Fc-region (Omalizumab) [15] that can be injected. Because of the high costs and the side effects, it is only used in severe cases.

A treatment of allergy in a causal way is the allergen-specific immunotherapy (SIT). In this procedure, patients get a low dose of allergen extracts or hypoallergenic constructs either injected (SCIT) or sublingual (SLIT) applicated that reduces the allergic symptoms in many cases [5].

1.2 Allergens

Allergens that cause Type-I hypersensitivity are proteins of different sizes and modifications that can be found in all kingdoms of life. Allergenic proteins can be inhaled, ingested, injected or touched and can lead to the typical symptoms depending on the site of exposure. Still today, evermore new allergens are investigated that can be grouped in many different superfamilies that in some parts contain still more homologues.

This work focuses on the pollen allergen Bet v 1 and its homologues (following chapters).

1.2.1 Bet v 1 and its homologues

Bet v 1 is the major allergen from the white birch tree *Betula pendula* and is the immunologically best studied allergen so far. It is estimated that 100 million people are sensitized to Bet v 1, especially in the Northern hemisphere (Northern America and Central Europe). 90% of all birch-pollen sensitized patients display specific IgE-antibodies for Bet v 1 which makes the birch pollen allergy one of the most common allergic disorders [16].

Bet v 1 is a member of ubiquitous pathogenesis-related protein family 10 (PR-10) with a high number of homologous proteins in all superkingdoms. Therefore, the Bet v 1-like family is evolutionally highly conserved [17]. Because of the widely spread occurrence of Bet v 1-related proteins in plants, allergic cross-reactions of Bet v 1 –sensitized patients to certain homologues are very common. Roughly 60% of birch pollinosis patients show allergic symptoms to certain vegetables and fruits [13, 18]. In the following chapters, Bet v 1 with its scaffold and conserved features is characterized as well as the phenomenon of Bet v 1-specific cross-reactivity to plant food homologues. Furthermore, Api g 1, the homologue from celery is introduced, the second important player concerning this work beside Bet v 1.

1.2.1.1 Bet v 1

Bet v 1 is a 17.5 kDa acidic protein with highly conserved fold [17]. So far, the structure could be determined in the crystalline state via X-ray crystallography as well as in solution by NMR-spectroscopy [19]. Figure 4 shows the three-dimensional structure of Bet v 1 and the topology diagram of the fold.

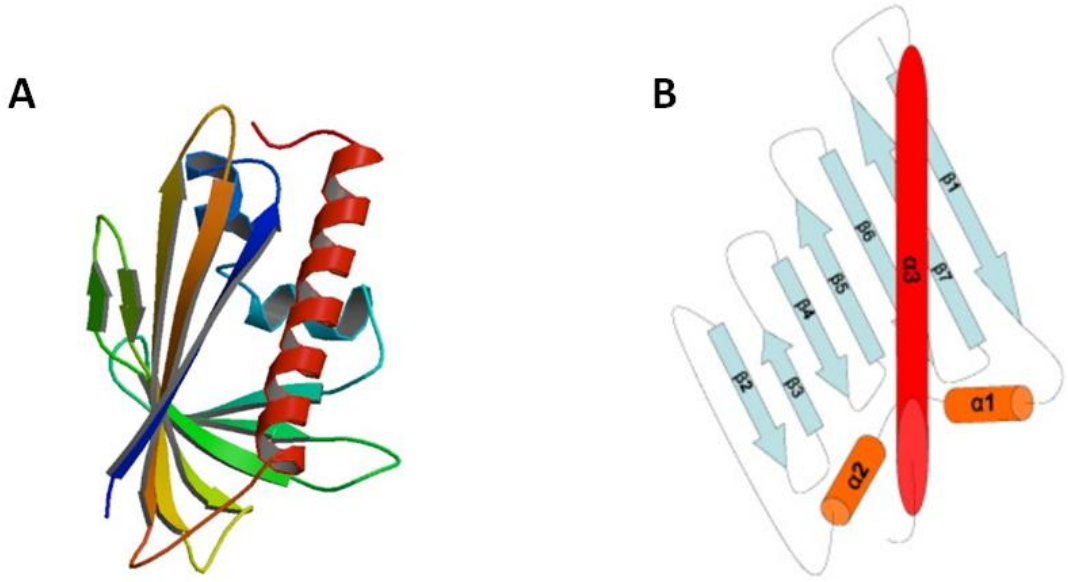


***Figure 4**: A: 3D- Structure of Bet v 1 (PDB code: 1BV1). B: schematic representation of the Bet v 1-like fold* **[20]**

The scaffold is characterized by a seven-stranded β-sheet and a large main α-helix that form a 30 Å-long cavity along the entire protein. This cleft is stabilized by two small α-helices that keep the main helix away from the β-sheets resulting in this long cavity. It has been shown that Bet v 1 can bind various ligands to its internal hydrophobic cleft. Crystallogrphically mapped ligand binding can reveal some key residues for ligand interaction and have shown that this process is very complex. It seems that Bet v 1 has a transport and storage function via the binding of ligands such as lipids, metabolites or phytohormones [20]. Studies have shown that this binding is highly dynamic and modulates the compactness of the protein [21]. The immunogenic impact of the ligand binding is still not fully understood.

1.2.1.2 Plant food homologues

As described before, Bet v 1 has a highly conserved fold which is especially very common in the kingdom of plants. Beside the existence of pollinosis protein homologues of Bet v 1, the homologous proteins found in certain plant foods play an important role in allergic cross-reactivity. People that are sensitized to birch pollen have a risk to react to the exposure on different fruits and vegetables upon ingestion. The reason for this effect is the recognition of specific conserved regions in these homologues by the Bet v 1-specific IgEs which leads to an allergic reaction with the characteristic symptoms, called the oral allergy syndrome. The list of Bet v 1 homologues represents a variety of plant foods (Table 1).

Table 1: *List of Bet v 1 homologues in certain plant foods [AllFam Database of allergen families, AF069:Bet v 1-related protein; October 2015]*

Allergen name	source	Allergen name	source
Act c 8	*Actinidia chinensis* (Gold kiwi)	**Gly m 4**	*Glycine max (soybean)*
Act d 11	*Actinidia deliciosa* (green kiwi)	**Lyc e 4**	*Lycopersicon esculentum (tomato)*
Act d 8	*Actinidia deliciosa* (green kiwi)	**Mal d 1**	*Malus domestica (apple)*
Api g 1	*Apium graveolens* (celery)	**Pru ar 1**	*Prunus armeniaca (Apricot)*
Ara h 8	*Arachis hypogaea* (peanut)	**Pru av 1**	*Prunus avium (sweet cherry)*
Cas s 1	*Castanea sativa (chestnut)*	**Pru p 1**	*Prunus persica (Peach)*
Cor a 1	*Corylus avellana (hazel)*	**Pyr c 1**	*Pyrus communis (pear)*
Dau c 1	*Daucus carota (carrot)*	**Rub i 1**	*Rubus idaeus (Raspberry)*
Fra a 1	*Fragaria ananassa (strawberry)*	**Vig r 1**	*Vigna radiata (mung bean)*

Beside the cross-allergic effect of IgEs, the Bet v 1-specific T cell epitopes play an important role in pollen-related food allergens by the activation of T-cell mediated late phase allergic reactions. These epitopes are not fully destroyed by food processing or digestion [13]. Thus, the impact of Bet v 1 homologues in plant foods is very high.

1.2.1.3 Api g 1

The Bet v 1-homolog Api g 1 is the major allergen from celery *Apium graveolens* [22]. As already described, IgE-reactivity and allergic symptoms to this food allergen are found in people that are primary sensitized by Bet v 1, so Api g 1 is a cross-reactive allergen [23]. This 17-kDa protein shows 64% similarity and 42% identity to Bet v 1 on the amino acid sequence level [22]. The IgE-reactivity tested by immunoblotting and skin-prick tests shows no significant difference between naturally extracted Api g 1 from celery and recombinantly produced Api g 1 (rApi g 1) [22, 23]. So far, the structure of Api g 1 could be solved by X-ray crystallography with a resolution of 2.9 Å (Fig. 5) [24]. In the crystal, Api g 1 shows a high structural similarity to its homolog Bet v 1. The major deviations are in the glycine-rich region known as the P-loop (residue 23-45) and in the loop regions between β3 and β4 as well as β5 and β6. The investigation of the molecular surface and sequential alignments reveal three conserved patches of the surface that seem to be responsible for the cross-reactivity with Bet v 1 [24]. In comparison with other Bet v 1-related food allergens, Api g 1 has a lower IgE-binding capacity and therefore, the cross-reactive allergenicity is weak [25].

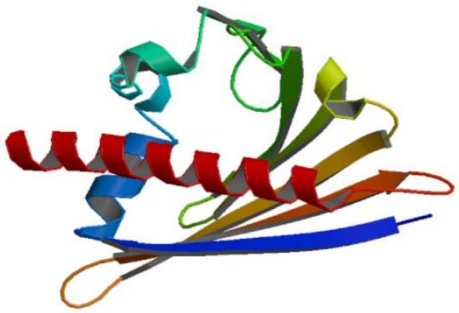


Figure 5: *Structure of Api g 1 (PDB-Code: 2BK0)*

In sum, the knowledge of the structural features as well as the differences in the IgE-binding of Api g 1 in comparison to Bet v 1 makes this allergen an important player in the chimeric construct designs concerning this work (following chapter).

1.3 The Api-Bet chimeric constructs

The mapping of conformation IgE-epitopes is crucial for the development of novel diagnostic and therapeutic approaches. One technique for this mapping is the design of chimeric constructs of the allergen of interest and a structural related homologue. Thereby, the areas of interest of an allergen are grafted onto the structure of a protein with low or no IgE-binding capacity. The corresponding surface-accessible residues are mutated yielding a chimeric construct with one area of the allergen of investigation onto another structurally related protein.

To map conformational IgE-epitopes of Bet v 1, this technique was applied by Barabara Gepp in the group of H. Breiteneder (Medical University of Vienna) using Api g 1 as the structurally related Bet v 1-homologue.

So four chimeric constructs (Api-Bet 1-4) were designed that enable the grafting of different surface areas of Bet v 1 on its low-allergenic counterpart Api g 1. Api g 1 has a high structural similarity, but the IgE-binding capacity is very low compared to other Bet v 1-homologs and furthermore, it is a non-sensitizing allergen. This establishes Api g 1 as a good basis for chimeric designs. In the chimeras, the sequence follows that of Api g 1 with 10-13 point mutations of the surface-accessible amino acid residues to the Bet v 1 counterpart (Fig. 6).

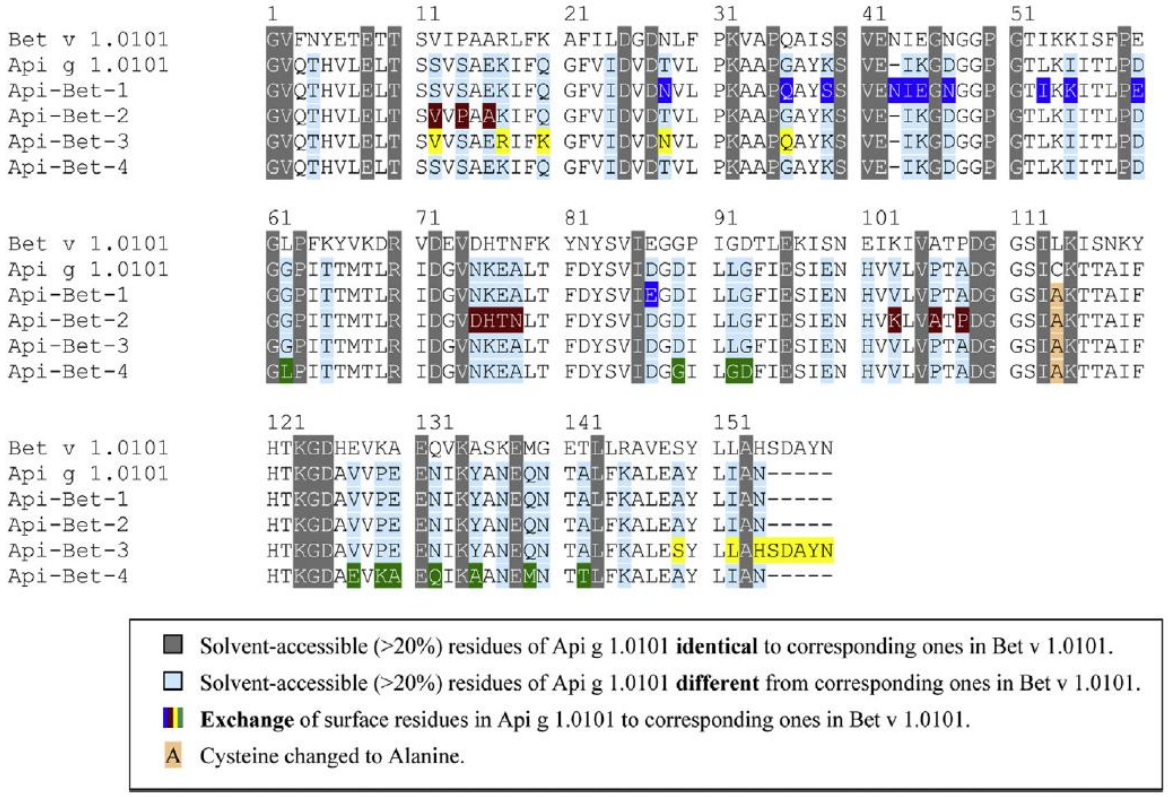


Figure 6: *Multiple sequence alignment of the four chimeras Api-Bet 1-4 and the parent proteins Bet v 1 and Api g 1* **[25]**.

Based on the scaffold of Api g 1, the residue exchanges in the four constructs were set in different quarters of the surface areas. In each quarter, the solvent-accessible residues that differ in both parent proteins were mutated to Bet v 1-specific residues. So in sum, nearly one fourth of the surface of each chimera equals that of the Bet v 1 corresponding area. In Api-Bet 1, the region known as the p-loop is mutated. In the next chimera, Api-Bet 2, the region opposite of this p-loop is replaced. In Api-Bet 3, the Bet v 1 -C-terminus and the surrounding region is grafted and the C-terminal α-helix is the mutated area in Api-Bet 4. The illustration of the mutated surface areas of the four chimeras is presented in Fig. 7. [25]

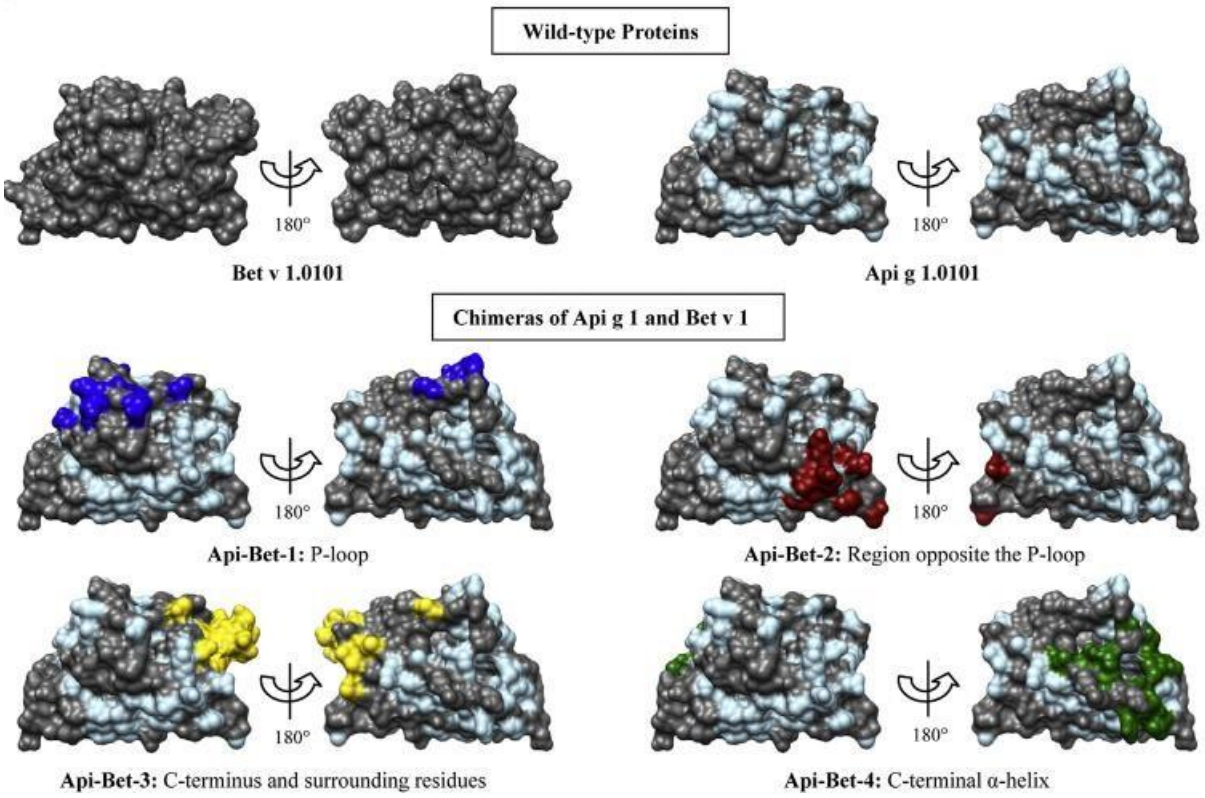


Figure 7: *Surface representation (front and back) of the four chimeric constructs and the parent proteins Api g1 and Bet v 1. The mutated surface residues are indicated by color* **[25]**.

All the constructs were designed based on the available crystal structures of Api g 1 and Bet v 1. The knowledge of the surface structure and the solvent-accessible amino acid residues enables the grafting of one area onto another structural equivalent part.

1.3.1 Api-Bet 1

In this chimeric construct, the p-loop region is the area of interest with replaced amino acid residues. This includes mutations in the loop regions, parts of the small α-helices and the adjacent β-sheet part (Fig. 8).

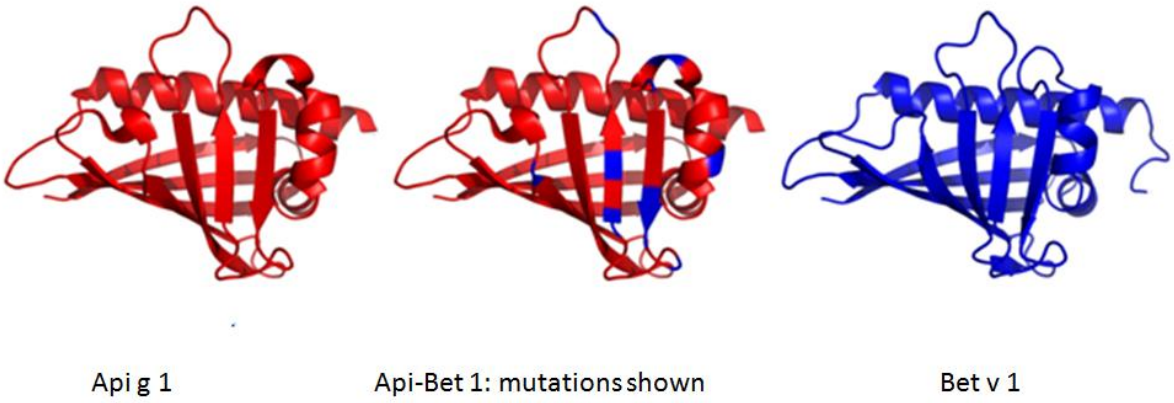


Figure 8: *Representation of the mutated areas of Api-Bet 1 with the expected structure and the comparison to the parental structures of Api g 1 and Bet v 1*

In sum, 10 point mutations were introduced including charged and uncharged amino acid residues.

1.3.2 Api-Bet 2

In Api-Bet 2, the area on the opposite of the p-loop is replaced. This results in 10 point mutations in the sequential part of residue 75-109 (Fig. 9).

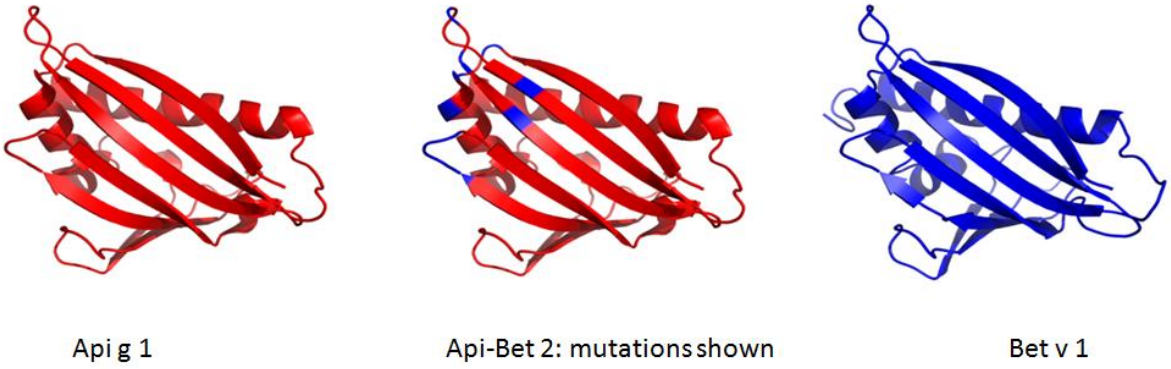


Figure 9: *Representation of the mutated areas of Api-Bet 2 with the expected structure and the comparison to the parental structures of Api g 1 and Bet v 1*

The mutated areas are in the loop regions as well as in parts of the β-sheets and in the first small α-helix.

1.3.3 Api-Bet 3

Api-Bet 3 is a construct with mutations at the C-terminus and the surrounding region (Fig. 10).

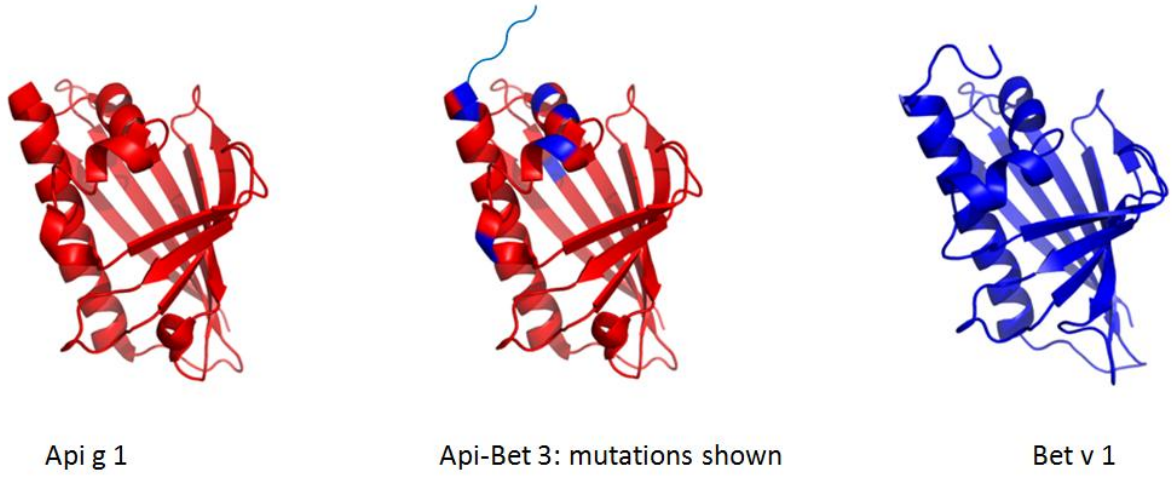


Figure 10: *Representation of the mutated areas of Api-Bet 3 with the expected structure and the comparison to the parental structures of Api g 1 and Bet v 1*

The terminal α-helix is elongated with five Bet v 1-residues resulting in an additional flexible tail like in its parent structure of Bet v 1.

1.3.4 Api-Bet 4

In the last chimera, Api-Bet 4, the surface-accessible residues of the main α-helix are exchanged (Fig. 11)

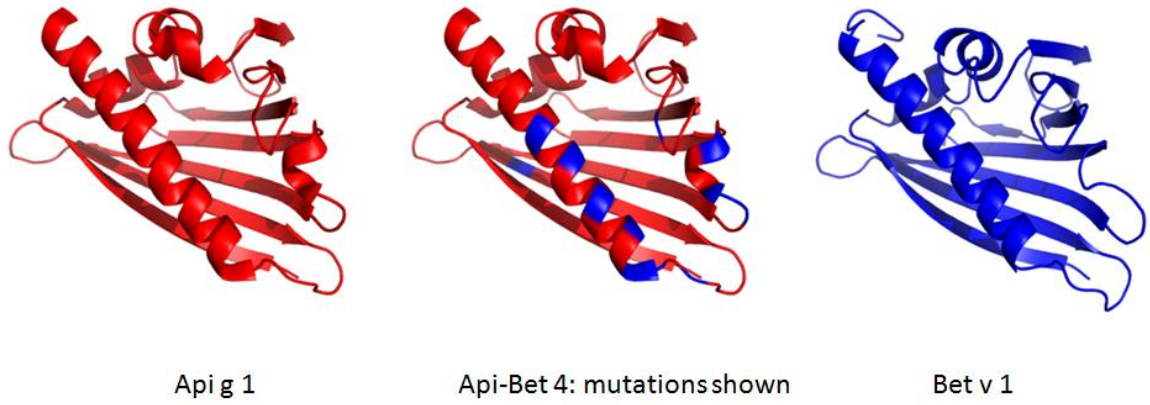


Figure 11: *Representation of the mutated areas of Api-Bet 4 with the expected structure and the comparison to the parental structures of Api g 1 and Bet v 1*

Additionally, this includes mutations of amino acid residues that are near on the surface to the C-terminus like parts of the small α-helix, the surrounding loop regions and the contacting β-sheet.

1.3.5 Api-Bets and IgE-binding

All of the four chimeras were already tested with the sera of Bet v 1-sensitized patients regarding their IgE-binding capacity using ELISA [25]. The specificity of these bindings was assessed by inhibition ELISAs. The Api-Bet constructs were recognized by 48%-89% of the patients. About 8% of the tested Patients did only recognize one chimera, whereas 31% showed an increased IgE-binding to all four chimeric constructs. Via these studies using the chimeric Api-Bet constructs it can be concluded that the Bet v 1-specific IgE-binding in patients is polyclonal and highly patient-specific. The IgE-epitopes of Bet v 1 are spread over the entire surface of Bet v 1 and the existence of one single major IgE-epitope can be excluded.

2 Aim of the master's thesis

The aim of this work was the structural and biophysical characterization of the four chimeric constructs of Bet v 1 and its homolog Api g 1 (Api-Bet 1-4). Thereby, the focus was the investigation of the impact the mutations have on the 3D-structure and the overall fold, as well as on the biophysical and biochemical properties of the proteins compared to the parental ones. The first part of the practical work was to establish expression and purification strategies yielding enough clean and properly folded protein for experimental approaches. The structural determination should be performed using X-ray crystallography.

3 Methods

In this chapter, the practical work and methods are described in detail. For the characterization of the four chimeras, small amounts of pre-cleaned protein from the cooperation group in Vienna were available, especially Api-Bet 1 and 3. Because in the structural biology, large amounts of properly folded, clean protein in a defined homogenous state are needed, the first strategy concerning this work was to establish the expression and purification of the chimeras for a further characterization.

3.1 Sequence Analysis

The primary sequences of all constructs were used for the calculation of biophysical and theoretical parameters that are important for further treatment and investigations, as well as for several measurements. The calculations were performed with the ProtParam tool of the ExPASy database [26]. Beside the amino acid residue composition, the molecular weight, the theoretical pI and the extinction coefficients can be calculated.

3.2 Tertiary structure prediction

The calculation of the most plausible theoretical tertiary structure is an important procedure to get information about the expected fold. Additionally, the calculated models can be used for further calculations and comparisons with measured data.. The tertiary structure prediction of the four chimeras was performed using Phyre² [27].

3.3 Cloning and Expression

All four Api-Bet constructs were sent in a pET 28a (+) expression vector with a kanamycin resistance gene. The genes were cloned in with EcoRI and NcoI restriction enzymes. In these expression systems, the proteins have no tag for purification or to increase the solubility. Because of a solubility problem of Api-Bet 2 and Api-Bet 3 during the expression in Vienna, these two chimeras were cloned into a pET 32a (+) expression fused to thioredoxin with an N-terminal His-tag. Between both protein sequences, a linker region with a PreScission Protease cleavage site enables the cleavage of the fusion construct (Fig. 12)

Figure 12: *Scheme of the Api-Bet -Thioredoxin fusion constructs*

Because the purification process of the Api-Bets without any affinity tag seemed to be very difficult and time consuming and correlates with a great loss of protein amount during the clean-up, the constructs should be cloned into another vector expression system with an N-terminal His-tag for a better purification using IMAC (the only available His-tag constructs were the THX-fusion constructs of Api-Bet 2 and 3). The vector of choice was the pJC 40 expression vector with a His-tag followed by a Factor Xa-cleavage site after which the target gene (api-bet 1-4) should be cloned in. A schematic representation of the protein construct is shown in Fig. 13.

Figure 13: *Scheme of the Api-Bet constructs with His-tag and Cleavage-site*

For the cloning, primers containing a BamHI and NdeI restriction sites were designed and subsequently used for a PCR. The target DNAs were the pET 28a (+) vectors containing the genes for the Api-Bet chimeras.

Forward Primer

GCGTTACATATGGGCGTACAGACC

Reverse Primer

CGGGATCCTTAATTCGCAATCAGATAGGCTTC

The PCR amplificates and the vector were digested with both restriction enzymes for 4h at 37°C in the following setup:

Table 2: *Set-up for Restriction enzyme digest of Api-Bet 1, 2 and 4 PCR-amplificates and the pJC 40 Vector DNA*

	Api-Bet 1	Api-Bet 2	Api-Bet 4	Vector
BamHI	1µL	1µL	1µL	1µL
NedI	1µL	1µL	1µL	1µL
CutSmart Buffer (10x)	5µL	5µl	5µL	5µL
Template DNA	30µL (100ng/µL)	30µL (100ng/µL)	30µL (100ng/µL)	30µL

ddH_2O	13µL	13µL	13µL	13µL
Total	50µL	50µL	50µL	50µL

After the cleaning of the cut genes and the vector by Agarose gel electrophoresis, the pJC 40 vector was dephosphorylated using arctic phosphatase (NEB) for 1h at 37°C. Afterwards the ligation was performed with T4-Ligase over night at 18°C.

***Table 3**: Set-up for the Ligation of the three api-bet genes into the pJC 40 vector*

	Api-Bet 1	Api-Bet 2	Api-Bet 4
T4-Ligase	1µL	1µL	1µL
T4-Ligase Buffer (10x)	2µL	2µl	2µL
Insert	7µL (0.0075pmol)	7µL (0.0075pmol)	7µL (0.0075pmol)
Vector	10µL 0.0025pmol	10µL 0.0025pmol	10µL 0.0025pmol
Total	20µL	20µL	20µL

The whole ligation batch was transformed into chemical competent Top10 *Escherichia coli* cloning cells using a protocol with heat shock for 45s at 42°C and plated on LB-Agar plates with Ampicillin and incubated over night at 37°C.

After a MiniPrep (QUIAGEN) of four colonies of each batch, the plasmids were purified and sent to a sequencing facility to proof the cloning success.

3.3.1 Transformation

All constructs were transformed into BL21 (DE3) *Escherichia coli* cells as the expression system. The transformations of all vectors were done with the same protocol for chemical transformation with heat shock. 10 µL of the vector (30-48 ng/µL) were added to 200 µL chemical competent cells on ice and incubated for 30 min. The heat shock was done for 45 seconds on 42°C in the water bath. After chilling on ice for 10 min, 1mL LB-Medium was added to each sample with a following incubation for 45 min 37°C. All cells (pelleted after short centrifugation) were plated on LB-Agar plates with Kanamycin or Ampicillin (depending on the expression vector) and incubated over night at 37°C.

3.3.2 Expression of Api-Bet 1,2 and 4 in pET 28a (+)

For the overnight cultures, a single colony of every construct (1, 2 and 4) was picked and inoculated into 200 mL sterile LB-medium with 50 µg/mL Kanamycin. The cells were incubated overnight at 37°C with 180 rpm. 20 mL of each ONC was used for inoculation of the main expression culture with 400 mL sterile LB-medium (with 50 µg/mL Kanamycin). The cells were grown to an OD_{600} of 0.6-0.8. For the induction of expression, IPTG was added to a final concentration of 1mM and the cultures were left shaking for 3.5 hours. The cells were harvested by centrifugation and the pellets were stored at -20°C.

3.3.3 Expression of Api-Bet 3-Thioredoxin in pET 32a (+)

For the ONC of the Api-Bet 3-Thioredoxin fusion construct, a single colony was inoculated into 200mL sterile LB-medium with 100 µg/mL Ampicillin and incubated overnight at 37°C with 180 rpm. 20mL of the overnight culture was used for the inoculation of the main culture (400 mL sterile LB-medium with Ampicillin). The cells were grown to an OD_{600} of 0.6 and then cooled down to 16°C. The induction was performed with 1mM IPTG (final concentration) and the cells were left shaking overnight (18 hours) at 16°C. The cells were harvested by centrifugation and the pellet was stored at -20°C.

3.3.4 Expression of Api-Bet 2 and 4 in pJC 40

The expression and ONC set-up of the two Api-Bet-constructs (2 and 4) with the N-terminal His-tag (pJC 40 expression vector) was the same as for the Api-Bet 3-Thioredoxin fusion construct (chapter 3.2.3.). Additionally to the cold expression at 16°C overnight, both proteins were also expressed for 4 hours at 37°C with the same amount of IPTG. The cells were harvested by centrifugation and the pellets were stored at -20°C.

The expression was controlled by SDS-PAGE with samples of the culture (500 µL) taken before induction and after expression.

3.4 Purification

In this chapter, the purification process from the cell pellets to clean protein is described. For the different constructs, Api-Bet 1, 2 and 4 without any purification tag, the Api-Bet 3-THX fusion construct and Api-Bet 2 and 4 with His-tag, different purification strategies were applied.

3.4.1 Api-Bet 1, 2 and 4 without tag

3.4.1.1 Lysis

The cell pellets of Api-Bet 1,2 and 4 were refrozen on ice and afterwards resuspended in 1*PBS. All batches were sonicated at 30% for 20 min in an ice-water bath. Afterwards, the lysed cells were centrifuged at 20 000g for 30 min (JA 300 rotor). The supernatant and the pellets were stored.

3.4.1.2 Refolding

The pellet containing the inclusion bodies, cell brackets and other insoluble materials was resuspended in 1*PBS and centrifuged at 16 000g for 20 min. The supernatant was discarded and the pellet was washed again in 1*PBS. This procedure was repeated once more to remove rests of soluble parts. Finally, the washed pellet containing the mostly clean inclusion bodies (IB) was solved in 3 mL of 8M Urea. After centrifugation at 16 000g for 20 min the supernatant, containing the denaturated protein was pooled.

The method of choice for the refolding of the protein constructs was rapid dilution. 50mL of cooled 1*PBS buffer was stirred and the unfolded protein in urea was added in 50µL steps. After the addition of the whole urea-protein sample, 1mL was taken and centrifugated full-speed. The supernatant was taken for protein concentration determination using NanoDrop. This process was performed for Api-Bet 1, 2 and 4.

To proof the success of the refolding, 10 µL of each batch after centrifugation and 10 µL of each urea sample were mixed with 10 µL 2*PGLB and taken for SDS-PAGE. The samples were cooked for 10 min and afterwards, 20µL of each sample and 5µL of a marker were loaded onto a 12% Tris/Glycin SDS-gel. The electrophoresis was run at 300V for 40 min. The staining was performed with Coomassie brilliand blue and the destaining was done in severe cycles with boiling water .

3.4.1.3 SEC

The refolded protein samples (50mL) were concentrated using Amicons (Merck) to a final concentration of about 2 mg/mL. After full-speed centrifugation for 20 min, 500 µL were loaded onto a Superdex 75 (GE Healthcare) column, respectively. The running buffer was 1*PBS with a flow of 0.5 mL/min. Fractionation was set to 4 mL, in the peak region, the fraction size was reduced to 0.5 mL. 10 µL samples of all peak-fractions as well as the loaded sample and the pellet after centrifugation and resuspension were mixed with 2*PGLB and loaded onto a 12%-Tris/Glycin SDS-Gel. The SDS-PAGE was run at 300 V for 30-40 min. The protein staining was done with Coomassie Brilliant blue.

3.4.2 Api-Bet 3-Thioredoxin fusion construct

3.4.2.1 Lysis

The lysis of the cell pellets from expression was performed in the same mode as for Api-Bet 1, 2 and 4 without tag using sonication.

3.4.2.2 Refolding

The refolding of the Api-Bet 3-THX protein construct was performed following the same strategy as for the tag-free constructs. The IBs were washed twice in 1*PBS and the solved in 8M urea. The rapid dilution was done in a 20mM Tris-buffer at pH 7.4 with 20% glycerol and 100mM NaCl. The whole batch was centrifuged full-speed, filtered (0.45 nm filter) and applied to an immobilized metal affinity chromatography (IMAC).

3.4.2.3 IMAC

The His-trap purification was performed with a 5mL Ni-NTa column. The full batch with the refolded protein was loaded onto the column with a flow of 4mL/min. The column was washed with 10 CV of the washing buffer and afterwards eluted with the elution buffer

Table 4: *Buffers for the IMAC-purification*

Buffer	Composition
Washing	20mM Tris (pH 7.4) 100mM NaCl 20% Glycerol
Elution	20mM Tris (pH 7.4) 100mM NaCl 300mM Imidazole 20% Glycerol

The eluate was collected in 1.5mL fractions and 10µl each were analyzed by SDS-PAGE.

3.4.3 Api-Bet 2 and 4 with His-tag

3.4.3.1 Lysis

The lysis of the cell pellets from expression was performed in the same mode as for Api-Bet 1, 2 and 4 without tag and Api-Bet 3-THX fusion construct using sonication.

3.4.3.2 Refolding

The refolding of the two His-tag constructs was performed using the same strategy as for the tag-free constructs and the Api-Bet 3-THX fusion construct. The IBs were washed twice in 1*PBS and solved in 8M urea. The rapid dilution was done in a 20mM Tris-buffer at pH 7.4 with 20% glycerol. The whole batch was centrifuged full-speed, filtered (0.45 nm filter) and applied to an immobilized metal affinity chromatography (IMAC).

3.4.3.3 IMAC

The His-trap purification was performed with a 5mL Ni-NTa column. The full batch with the refolded protein was loaded onto the column with a flow of 4mL/min. The column was washed with 10 CV of the washing buffer and afterwards eluted with the elution buffer

Table 5: *Buffers for the IMAC-purification*

Buffer	Composition
Washing	20mM Tris (pH 7.4) 100mM NaCl 20% Glycerol
Elution	20mM Tris (pH 7.4) 100mM NaCl 300mM Imidazole 20% Glycerol

The eluate was collected in 1.5mL fractions and 10µl each were analyzed by SDS-PAGE.

3.5 Circular Dichroism analysis

Circular Dichroism (CD-) spectroscopy is a widely used method for analyzing chiral molecules, especially proteins. The basic principle of this method is the interaction of polarized light with these chiral chromophores resulting in an extinction difference between right- and left-circularly polarized light. The wavelength dependent measurement of this difference yields a characteristic CD-spectrum. The optically active groups in proteins are the aromatic side chains and, rather more important for this method because of the chiral compound, the polypeptide backbone, consisting of the amide bonds. The difference in exhibiting of left- and right-polarized light depends on the arrangement and the orientation of the polypeptide backbone in well-ordered structures, resulting in characteristic absorption spectra for each secondary structure element, such as α-helix, β-sheets, turns and

random coil-structures with typical maxima and minima of absorption. So, this method allows the determination of the overall structural composition of proteins in solution [28].
In comparison to high resolution methods like X-ray crystallography, no three-dimensional structure can be explored by Circular Dichroism analysis, but the overall order and the percentage composition of secondary structure elements can be analyzed. In the field of biophysical structure analysis of proteins, CD-spectroscopy is a quick method with little preparation needed and small amounts of sample to be used.

Beside the secondary structure determination, Circular Dichroism is used to investigate the stability of proteins against thermal and chemical denaturation effects and the influence of ligand binding on the structural integrity. For this procedure, the changes in the spectra are observed. For a melting analysis, the sample is heated and the changes in the full spectra can be observed. For tracing the process, the signal shift at a characteristic wavelength is detected and as a result, a melting curve is measured [28].

3.5.1 Full spectrum measurement

All four Api-Bet constructs were measured in 0.02 cm cuvettes in a jasco J-715 CD-spectrometer in a UV-range of 260-190nm. The full spectra were measured with 100 nm/min and a data pitch of 0.2 nm. The four samples were measured in the following setup:

Table 6: *Protein solutions with concentration and buffer used for the CD-measurements*

Protein construct	Volume [µL]	Concentration [mg/mL]	Buffer
Api-Bet 1	200	0.5	1*PBS
Api-Bet 2	200	0.26	1*PBS
Api-Bet 3	200	0.56	10 mM Na_3PO_4
Api-Bet 4	200	0.28	1*PBS

The CD-signal was later on transformed into mean residue epllipticity (MRE) that includes the protein concentrations, molecular weights and the number of peptide bonds.

3.5.1.1 Secondary structure determination

The full spectra were used for the calculation of the secondary structure element composition using Dichroweb [29]. For this calculation, the CDSSTR method [30] was used with dataset 4 and 7 [31].

3.5.2 Temperature Scans

The heat stability and melting behavior of the constructs were checked by temperature CD-scans. The samples were heated from 25°C to 95°C in 1°/min in a data pitch of 0.1 °C and measured at 208nm resulting in a melting curve. Every 5°, a full spectrum was measured with the same setup as described in chapter 3.5.1. Immediately after each Up-scan, a Down-scan from 95°C to the starting temperature of 25°C was performed in the same mode (1°/min, 208nm). These scans were also implemented as stepscans.

3.5.3 Refolding capacity

The ability of the four constructs to refold after heating and cooling down was checked by overlaying of the full spectra at RT before each temperature scan and the full spectra after the Down-scans to 25°C. A shift in the overlaid spectra indicates an incomplete refolding.

3.6 Crystallization

All Api-Bet constructs were concentrated to a final concentration of 5mg/mL. 0.5µL of the protein solutions were mixed with 0.5µL of the conditions of the crystallization screen a 96 well microbatch plates and covered with a paraffin/silicon oil (3:1) mixture. The following crystallization screens were used:

Table 7: *Proteins and Screens used for the microbatch vrystallization*

Protein	Crystallization screen
Api-Bet 1	Index JCSG+ PEG/ION
Api-Bet 2	Index JCSG+
Api-Bet 3	Index JCSG+
Api-Bet 4	Index JCSG+
Api-Bet 3-THX	Index Proplex

Optimization was performed of the PEG/ION condition 32:

0.2M Mg_2SO_4 20% PEG 3,350 pH 6.0

In the crystallization optimization, the pH was increased to 6.8, and the PEG concentration to 25%. The concentration of Mg_2SO_4 was varied from 50 to 200mM. The set-up was performed as a sitting-drop experiment with 5µL protein solution and 5µL of the reservoir solution.

3.7 X-Ray analysis

The crystals were primarily tested at the home source X-ray anode for diffraction. Thereby, possible salt crystals can be excluded from further optimizations and data collections. The best diffracting protein crystals were sent and measured at the European Synchrotron Radiation Facility (ESRF) in Grenoble for data collection.

3.8 Molecular Replacement

Molecular Replacement (MR) was the method of choice for solving the phase problem. Here, the structures of the parent proteins Api g 1 and Bet v 1 were used to get a refineable solution. The MR process was performed in the Phenix package (Python-based Hierarchical ENvironment for Integrated Xtallography) [32].

3.9 Structural Refinement

The diffraction data set was integrated with iMosflm [33] and merged and scaled using Scala, part of the ccp4 program package [34]. The structural refinement of the structure was performed with the Phenix program package [32] with connection to coot [35] for manual adaption. Molecular graphics of the protein structures were performed using PyMOL (The PyMOL Molecular Graphics System, Version 1.7.1 Schrödinger, LLC.) .

3.10 SAXS meassurment

Small angle X-ray scattering is a method that uses the scattering of X-rays in very low angles of about 0.1-1° around the beam stop. Another difference to the X-ray crystallography is that the protein

samples are measured in solution. With this method, a low resolution shape of the protein in solution can be determined and, especially important in this case, it reveals information about the oigomerization state of the measured protein [36].

Here, Api-Bet 1 was concentrated to a final concentration of 12.7 mg/mL and measured at the SAXS-beamline at the ESRF by Christian Fercher, University of Graz. The sample was diluted twice and the experiments were performed for each concentration.

Table 8: *Protein samples and concentrations for the SAXS experiment.*

Protein sample	Concentration [mg/mL]	Buffer
Api-Bet 1 (undiluted)	12.7	1*PBS
Api-Bet 1 (1. dilution)	6.2	1*PBS
Api-Bet1 (2. Dilution)	2.9	1*PBS

The data was processed using the ATSAS package [37]

4 Results

4.1 Sequence Analysis

4.1.1 Api-Bet 1

```
         10          20          30          40          50          60
GVQTHVLELT SSVSAEKIFQ GFVIDVDNVL PKAAPQAYSS VENIEGNGGP GTIKKITLPE

         70          80          90         100         110         120
GGPITTMTLR IDGVNKEALT FDYSVIEGDI LLGFIESIEN HVVLVPTADG GSIAKTTAIF

        130         140         150
HTKGDAVVPE ENIKYANEQN TALFKALEAY LIAN
```

Number of amino acids: 154

Molecular weight: 16357.5

Theoretical pI: 4.59

Amino acid composition:

Ala (A)	14	9.1%	
Arg (R)	1	0.6%	
Asn (N)	9	5.8%	
Asp (D)	7	4.5%	
Cys (C)	0	0.0%	
Gln (Q)	4	2.6%	
Glu (E)	13	8.4%	
Gly (G)	14	9.1%	
His (H)	3	1.9%	
Ile (I)	15	9.7%	
Leu (L)	12	7.8%	
Lys (K)	9	5.8%	
Met (M)	1	0.6%	
Phe (F)	6	3.9%	
Pro (P)	7	4.5%	
Ser (S)	8	5.2%	
Thr (T)	13	8.4%	
Trp (W)	0	0.0%	
Tyr (Y)	4	2.6%	
Val (V)	14	9.1%	

Total number of negatively charged residues (Asp + Glu): 20
Total number of positively charged residues (Arg + Lys): 10

Atomic composition:

Carbon	C	735
Hydrogen	H	1169
Nitrogen	N	185
Oxygen	O	233
Sulfur	S	1

Formula: $C_{735}H_{1169}N_{185}O_{233}S_1$
Total number of atoms: 2323

Ext. coefficient 5960
Abs 0.1% (=1 g/l) 0.364

4.1.2 Api-Bet 2

```
         10          20          30          40          50          60
MGVQTHVLEL TSVVPAAKIF QGFVIDVDTV LPKAAPGAYK SVEIKGDGGP GTLKIITLPD

         70          80          90         100         110         120
GGPITTMTLR IDGVDHTNLT FDYSVIDGDI LLGFIESIEN HVKLVATPDG GSIAKTTAIF

        130         140         150
HTKGDAVVPE ENIKYANEQN TALFKALEAY LIAN
```

Number of amino acids: 154

Molecular weight: 16306.6

Theoretical pI: 4.79

Amino acid composition:

Ala (A)	14 9.1%	Leu (L)	13 8.4%
Arg (R)	1 0.6%	Lys (K)	10 6.5%
Asn (N)	6 3.9%	Met (M)	2 1.3%
Asp (D)	11 7.1%	Phe (F)	6 3.9%
Cys (C)	0 0.0%	Pro (P)	8 5.2%
Gln (Q)	3 1.9%	Ser (S)	5 3.2%
Glu (E)	8 5.2%	Thr (T)	15 9.7%
Gly (G)	15 9.7%	Trp (W)	0 0.0%
His (H)	4 2.6%	Tyr (Y)	4 2.6%
Ile (I)	15 9.7%	Val (V)	14 9.1%

Total number of negatively charged residues (Asp + Glu): 19
Total number of positively charged residues (Arg + Lys): 11

Atomic composition:

Carbon	C	738
Hydrogen	H	1176
Nitrogen	N	185
Oxygen	O	226
Sulfur	S	2

Formula: $C_{736}H_{1176}N_{185}O_{226}S_2$
Total number of atoms: 2326

Ext. coefficient 5960
Abs 0.1% (=1 g/l) 0.365

4.1.3 Api-Bet 3

```
         10          20          30          40          50          60
GPMGVQTHVL ELTSVVSAER IFKGFVIDVD NVLPKAAPQA YKSVEIKGDG GPGTLKIITL

         70          80          90         100         110         120
PDGGPITTMT LRIDGVNKEA LTFDYSVIDG DILLGFIESI ENHVVLVPTA DGGSIAKTTA

        130         140         150         160
IFHTKGDAVV PEENIKYANE QNTALFKALE SYLLAHSDAY N
```

Number of amino acids: 161

Molecular weight: 17156.5

Theoretical pI: 4.75

Amino acid composition:

Ala (A)	14 8.7%	Leu (L)	14 8.7%
Arg (R)	2 1.2%	Lys (K)	10 6.2%
Asn (N)	7 4.3%	Met (M)	2 1.2%
Asp (D)	11 6.8%	Phe (F)	6 3.7%
Cys (C)	0 0.0%	Pro (P)	8 5.0%
Gln (Q)	3 1.9%	Ser (S)	8 5.0%
Glu (E)	10 6.2%	Thr (T)	13 8.1%
Gly (G)	15 9.3%	Trp (W)	0 0.0%
His (H)	4 2.5%	Tyr (Y)	5 3.1%
Ile (I)	14 8.7%	Val (V)	15 9.3%

Total number of negatively charged residues (Asp + Glu): 21
Total number of positively charged residues (Arg + Lys): 12

Atomic composition:

Carbon	C	773
Hydrogen	H	1227
Nitrogen	N	195
Oxygen	O	240
Sulfur	S	2

Formula: $C_{773}H_{1227}N_{195}O_{240}S_2$
Total number of atoms: 2437

Ext. coefficient 7450
Abs 0.1% (=1 g/l) 0.434

4.1.4 Api-Bet 4

```
         10          20          30          40          50          60
GVQTHVLELT SSVSAEKIFQ GFVIDVDTVL PKAAPGAYKS VEIKGDGGPG TLKIITLPDG

         70          80          90         100         110         120
LPITTMTLRI DGVNKEALTF DYSVIDGGIL GDFIESIENH VVLVPTADGG SIAKTTAIFH

        130         140         150
TKGDAEVKAE QIKAANEMNT TLFKALEAYL IAN
```

Number of amino acids: 153

Molecular weight: 16115.4

Theoretical pI: 4.74

Amino acid composition:

Ala (A)	15	9.8%	
Arg (R)	1	0.6%	
Asn (N)	5	3.3%	
Asp (D)	10	6.5%	
Cys (C)	0	0.0%	
Gln (Q)	3	1.9%	
Glu (E)	10	6.5%	
Gly (G)	15	9.8%	
His (H)	3	2.0%	
Ile (I)	15	9.8%	
Leu (L)	13	8.5%	
Lys (K)	11	7.2%	
Met (M)	2	1.3%	
Phe (F)	6	3.9%	
Pro (P)	6	3.9%	
Ser (S)	7	4.6%	
Thr (T)	15	9.8%	
Trp (W)	0	0.0%	
Tyr (Y)	3	2.0%	
Val (V)	13	8.5%	

Total number of negatively charged residues (Asp + Glu): 20
Total number of positively charged residues (Arg + Lys): 12

Atomic composition:

Carbon	C	725
Hydrogen	H	1167
Nitrogen	N	185
Oxygen	O	227
Sulfur	S	2

Formula: $C_{725}H_{1167}N_{185}O_{227}S_2$
Total number of atoms: 2302

Ext. coefficient 4470
Abs 0.1% (=1 g/l) 0.277

4.2 Tertiary structure prediction

The prediction results of the Phyre² server revealed the expected structures of the parent proteins for all four chimeric constructs. The model template for all chimeras was d2bk0a, the structure of the parent protein Api g 1. The Confidence and Coverage results are listed in the table below.

Table 9: *Results of the tertiary structure predction by ths Phyre² server.*

Chimeric onstruct	Confidence [%]	Coverage [%]
Api-Bet 1	100.0	99
Api-Bet 2	100.0	99
Api-Bet 3	100.0	99
Api-Bet 4	100.0	100

Because of the high sequence identity with Api g 1, the confidence of the models is 100% in all structure predictions. The coverage gives the percentage of the residues that could be modeled by the template structure. The modeled structures of the four constructs (Fig. 14) show the expected TBP-like fold of the PR-10-like family that all members of the Bet v 1-like superfamily are sharing. Interestingly, the second small α-helix is not well defined in the model of Api-Bet 3.

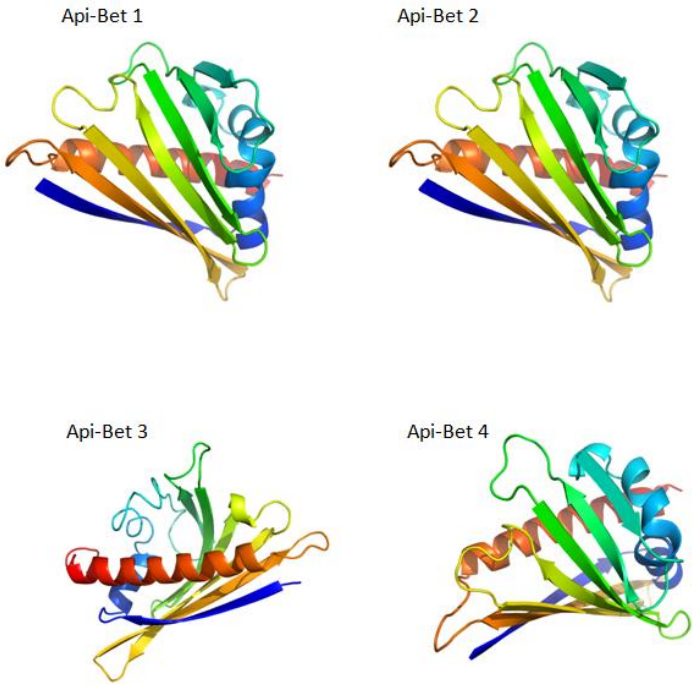
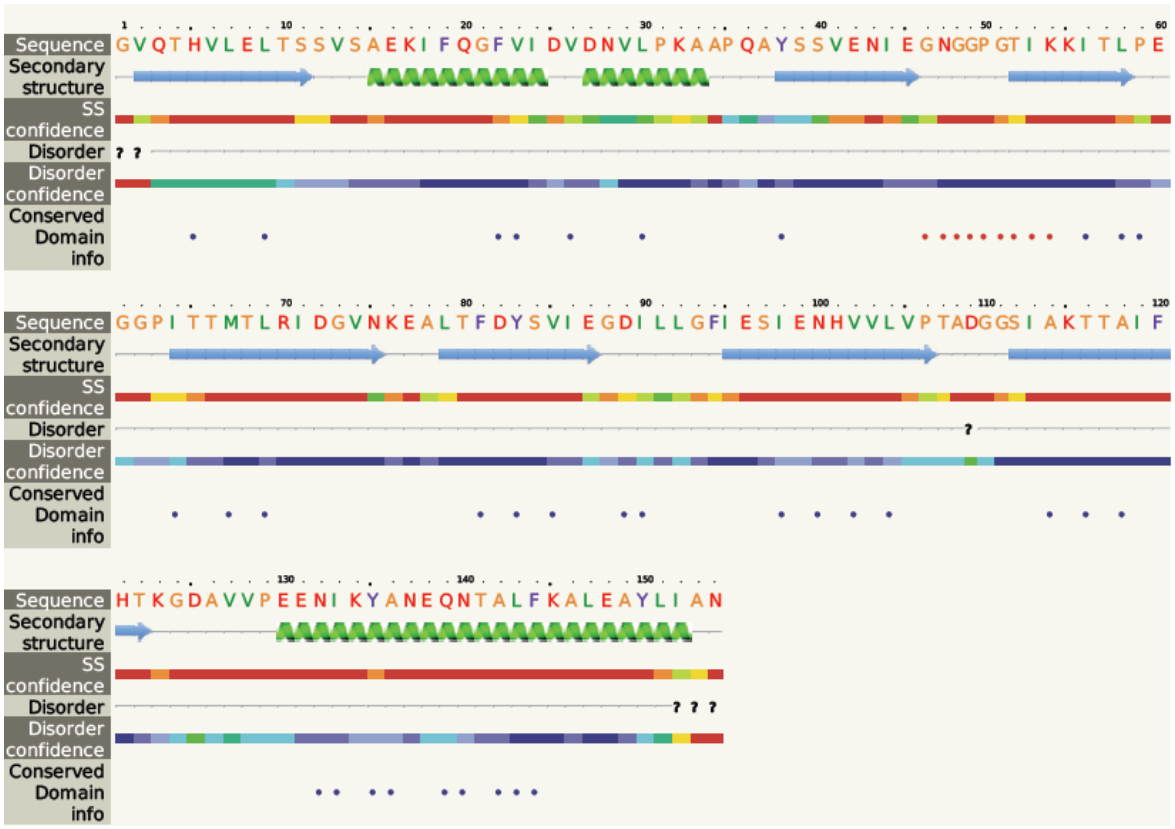


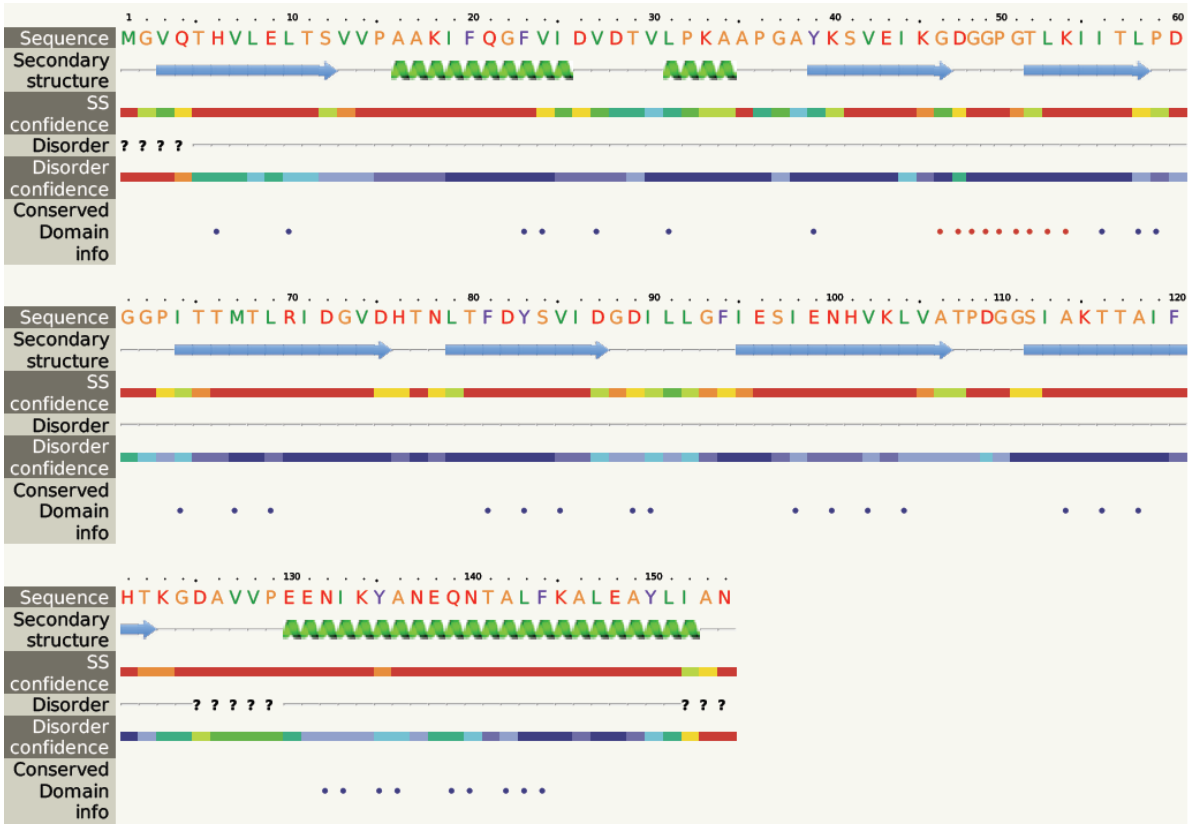
Figure 14: *Phyre²-modeled structures of the four Api-Bet constructs.*

The following data shows the detailed secondary structure and disorder prediction aligned on the sequence.

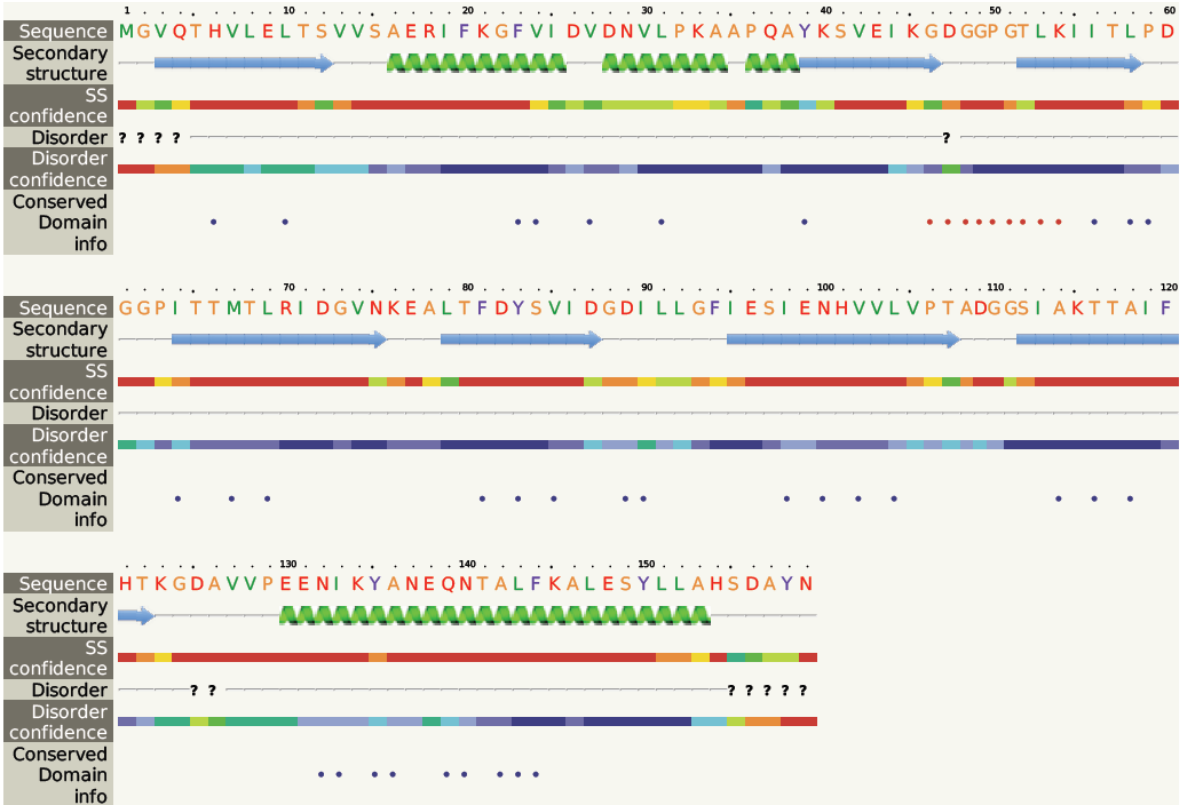
Api-Bet 1



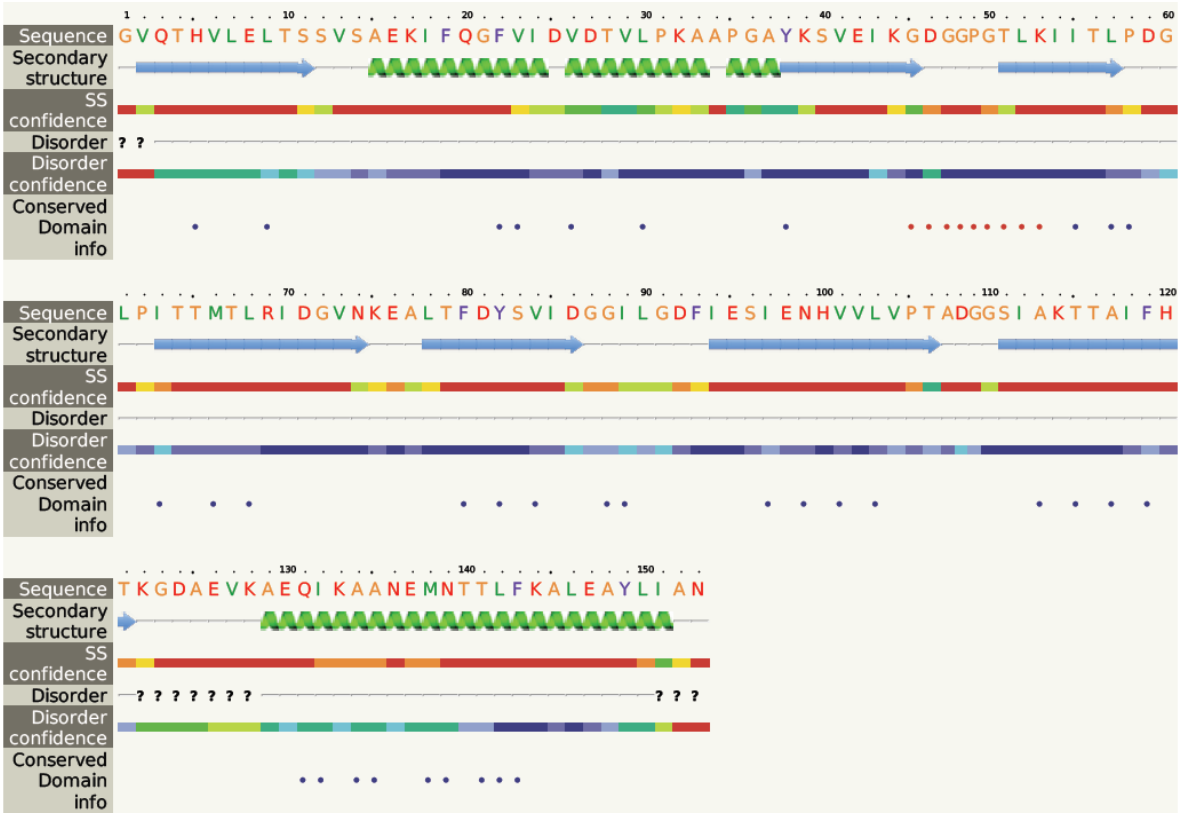
Api-Bet 2



Api-Bet 3



Api-Bet 4



4.3 Cloning and Expression

The cloning of the api-bet genes into the pJC 40 expression vector containing an N-terminal His-tag for purification was only successful for Api-Bet 2 and 4. Api-Bet 1 shows a lot of unspecific side products even after the PCR and the ligation was not successful after various trials.

4.3.1 Expression of Api-Bet 1, 2 and 4 without tag

The expression of Api-Bet 1, 2 and 4 without tag from the pET 28a (+) expression vector was successful. The SDS-PAGE (Fig. 15) shows the E. coli lysate before (left lanes) and after the induction of expression (right lanes).

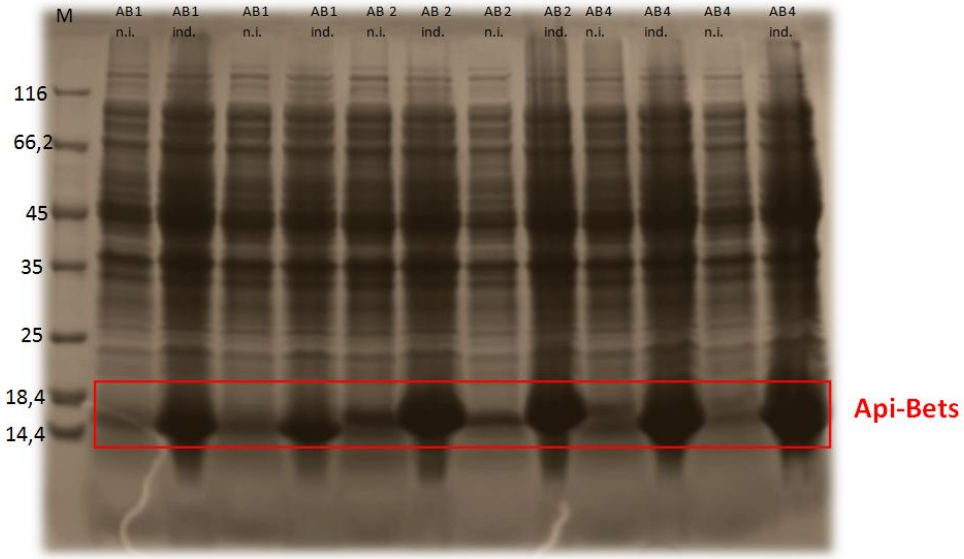


***Figure 15**: SDS-PAGE of the expression control of Api-Bet 1, 2 and 4. In the left slots, the samples before induction (n.i.) are shown and in the right slots the samples after induction for 3h (ind.) were loaded.*

All three Api-Bet constructs were highly overexpressed after 3h of induction with IPTG. Further characterization after the lysis of the cells revealed that nearly 60-70% of the protein is insoluble in inclusion bodies.

4.3.2 Expression of Api-Bet 3-Thioredoxin fusion construct

Also the expression of the Api-Bet 3-THX fusion construct from the pET 32a (+) expression vector was successful leading to large bands of overexpression at the right size on the SDS-Gel (Fig. 16)

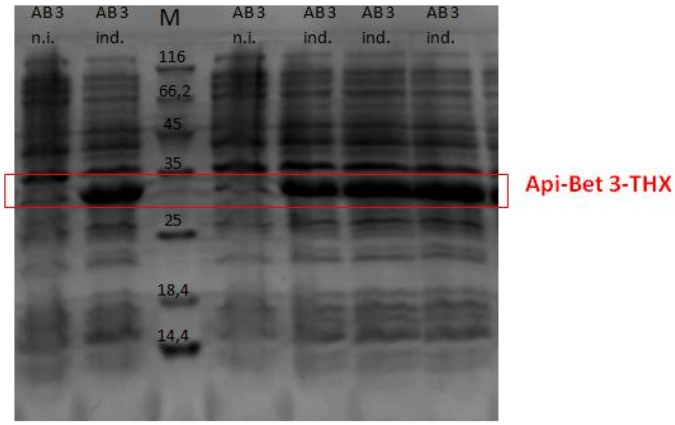


Figure 16: *SDS-PAGE of the expression control of Api-Bet 3-THX. The samples before induction (n.i.) and after induction overnight (ind.) are shown for the control of the expression success.*

After the lysis of the cells, about 80% of the protein constructs are in the insoluble pellet fraction, whereas only 20% were soluble. For the purification process, the insoluble fraction was taken for an inclusion body purification and refolding.

4.3.3 Expression of Api-Bet 2 and 4-His-tag variants

Api-Bet 2 and 4 with His-tag from the pJC 40 expression vector were nicely overexpressed after induction with IPTG overnight at 18°C. The expression control is sown on Fig. 17 and 18.

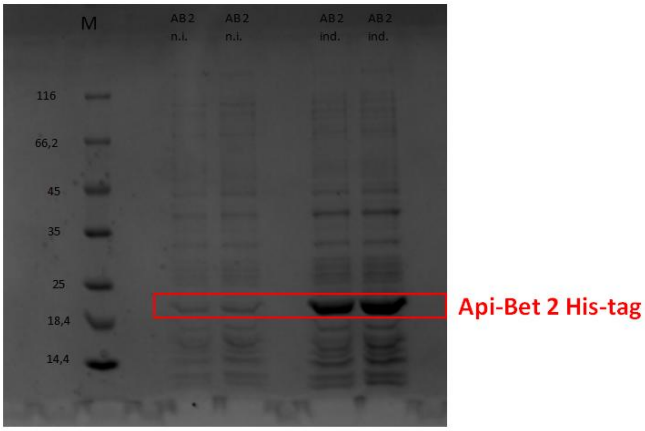


Figure 17: *SDS-PAGE of the expression control of Api-Bet 2 with His-tag. For the control the samples before induction (n.i.) and after induction and expression overnight (ind.) were loaded.*

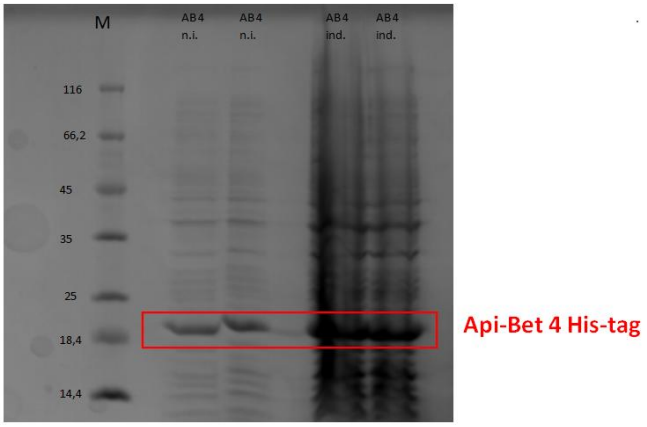


Figure 18: *SDS-PAGE of the expression control of Api-Bet 4 with His-tag. For the control the samples before induction (n.i.) and after induction and expression overnight (ind.) were loaded.*

Api-Bet 4-His shows a leaky expression even before induction with IPTG. Although both constructs, Api-Bet 2 and 4, show a high overexpression. The further characterization after cell lysis revealed the both protein constructs are completely insoluble. Therefore, the following protein purification was performed by IB-purification and protein refolding.

4.4 Purification

All different protein variants were purified by an inclusion body clean-up and refolding of the proteins followed by chromatographic techniques. The purification of the soluble fraction after expression was tried before and was very resource and time-consuming and correlated with high loss of protein during a variety of chromatographies, salt precipitations and dialyzing steps. Therefore, the purification of the larger insoluble protein fraction by refolding was tried.

4.4.1 Api-Bet 1, 2 and 4 without tag

4.4.1.1 Refolding

The refolding by rapid dilution of the three chimeras after the IB purification with 8M urea was successful. The control was performed by SDS-PAGE with the pellet fractions and the 50mL batches containing the diluted and refolded protein after centrifugation to get rid of possibly precipitated protein (Fig.19).

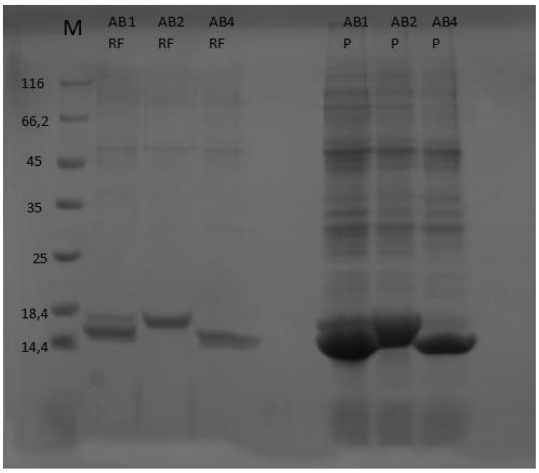


Figure 19: *SDS-PAGE of the pellet fractions (P) of Api-Bet 1, 2 and 4 containing the IBs with protein (the whole pellet solved in 3mL) in the right lanes and the refolded proteins (RF) in 50mL buffer in the left lanes.*

The pellets with the IBs had a high content of the target proteins and less impurities than a full lysate. The refolding by rapid dilution was successful in evidence by defined bands on the gel. Additionally, the amount of other proteins could be reduced.

4.4.1.2 Size exclusion chromatography

To get rid of the impurities and to separate the monomeric Api-Bet proteins from possibly aggregated or oligomerized proteins, a size exclusion chromatography was performed for each batch.

4.4.1.2.1 Api-Bet 1

Before the loading of the concentrated protein onto the column, the sample was centrifuged at 16000g for 15min. The chromatogram of the run is plotted in Fig. 20.

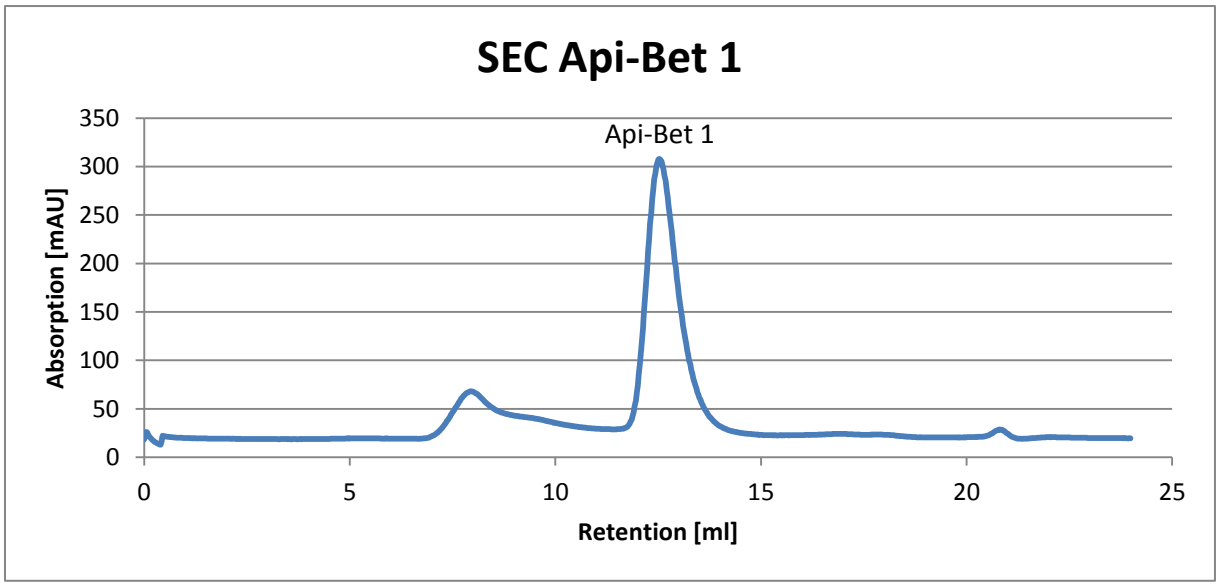


Figure 20: *Size exclusion chromatogram of Api-Bet 1 with a Superdex 75 column and a detection at 280nm.*

The chromatogram shows one main peak at a retention of 12-14mL and an aggregation peak at 7-8mL. All fractions of the peak areas as well as the loaded sample were investigated by SDS-PAGE to identify the proteins (Fig. 21).

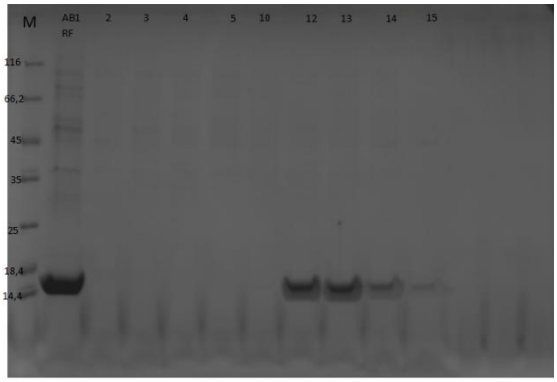


Figure 21: *SDS-PAGE of the SEC fractions yielding clean protein in the main peak (fractions 12-14) and the loaded protein sample (first lane)*

The SDS-PAGE shows that the main peak contents of Api-Bet 1 in a clean state. The impurities are found in the aggregation peak (fractions 2-5). The chromatogram shows no protein peak at a size for a dimeric protein. Because the IBs with Api-Bet 1 were washed two times more than the other protein batches and because it was higher concentrated than Api-Bet 2 and 4, the chromatogram shows a weak aggregation peak containing impurities in relation to the main peak.

4.4.1.2.2 Api-Bet 2

Before the SEC run, the Api-Bet 2 protein sample was centrifuged at 16000g for 15min yielding a small visible pellet. The supernatant was loaded onto the column and the chromatography was performed with the same setup as for Api-Bet 1. The chromatogram is shown in Fig. 22.

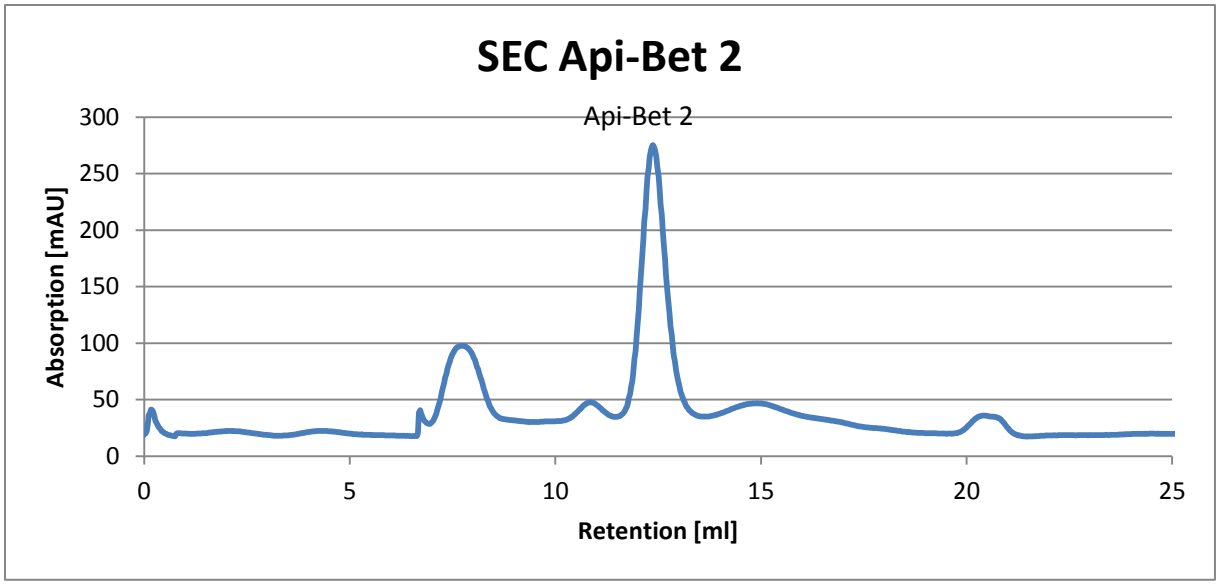


Figure 22: *Size exclusion chromatogram of Api-Bet 2 with a Superdex 75 column and a detection at 280nm.*

Because the protein sample has a higher content of impurities, the aggregation peak (retention of 7-8mL) is much larger than in the chromatogram of Api-Bet 1. A main protein peak is found again at a retention of 12-14mL. All fractions as well as the loaded protein sample and the pellet after centrifugation in PGLB and water (10+10µL) were used for SDS-PAGE. The result is shown in Fig. 23.

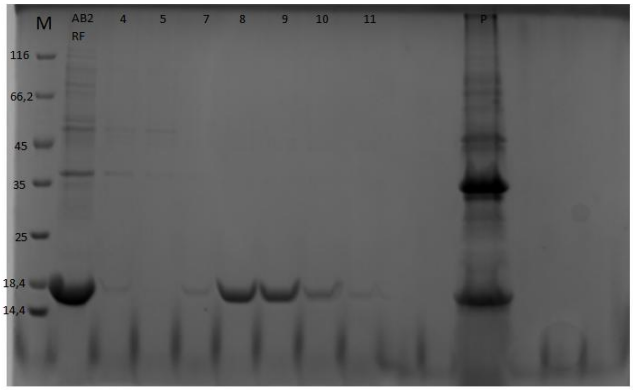


Figure 23: *SDS-PAGE of the SEC fractions yielding clean protein in the main peak (fractions 7-10) and the loaded protein sample (first lane). Additionally, the pellet after centrifugation was loaded (right lane).*

The pellet after centrifugation (last slot) revealed that a large amount of the impurities were insoluble and aggregated and could be separated by centrifugation. The remaining other proteins were found in the aggregation peak in the chromatogram (fraction 4-5). The SEC yielded clean Api-Bet 2 found in the fractions 7-10 that were pooled and stored at 4°C.

4.4.1.2.3 Api-Bet 4

The sample containing Api-Bet 4 was centrifuged at 16000g for 15min yielding a small pellet. The supernatant was loaded onto the Superdex 75 column and the run was performed as for the other samples. The chromatogram is plotted in Fig. 24.

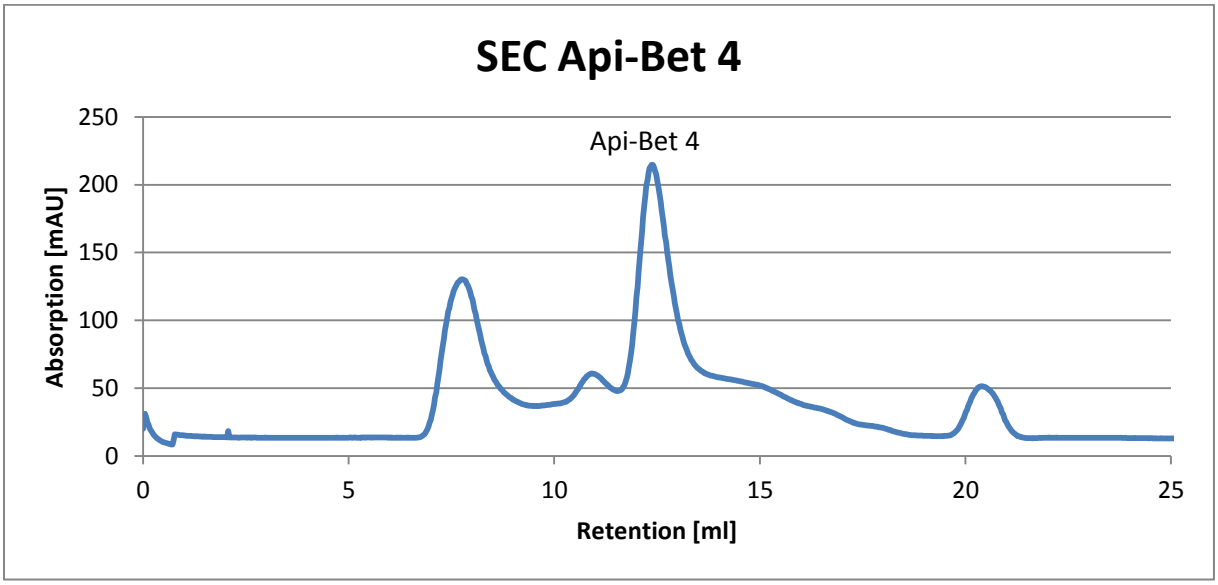


Figure 24: *Size exclusion chromatogram of Api-Bet 2 with a Superdex 75 column and a detection at 280nm.*

The chromatogram again shows a main peak at the same retention as for the other Api-Bets (12-14mL). Again, there is a high aggregation peak found and a small peak directly before the main peak and, at a retention of 20mL, another small peak can be found probably containing degraded protein fragments. All protein containing fractions as well as the loaded sample and the pellet after centrifugation were taken for SDS-PAGE shown in Fig.25.

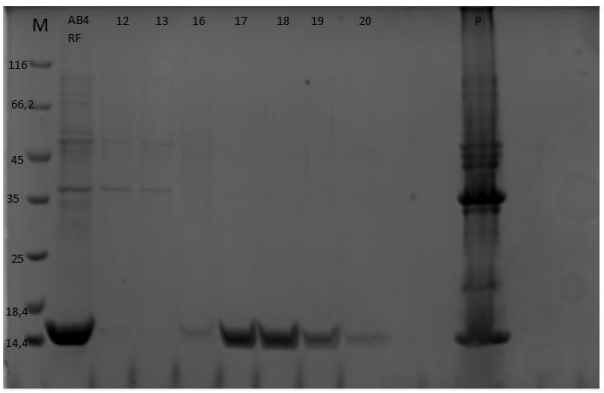


Figure 25: *SDS-PAGE of the SEC fractions yielding clean protein in the main peak (fractions 16-20) and the impurities in the aggregation peak (fractions 12-13). Additionally, the loaded protein sample (first lane) and the pellet after centrifugation was loaded (right lane).*

Again the SEC reveals clean protein separated from the impurities at the same retention. The pellet loaded onto the gel indicates that a large amount of the impurities precipitated after the refolding

and could therefore be separated. The fractions containing clean Api-Bet 4 were pooled and stored at 4°C.

4.4.2 Api-Bet 3-Thioredoxin fusion construct

4.4.2.1 Refolding

The refolding of the large amount of insoluble Api-Bet 3-THX from the IBs was successful. The protein concentration after rapid dilution in 50mL buffer could be determined at 0.3mg/mL. The whole batch was taken for IMAC purification.

4.4.2.2 IMAC

The IMAC was performed as described in chapter (3.4.2.3). All fractions, the loaded sample and the flowthrough during the washing steps (2 fractions) were taken for SDS-PAGE. The result is shown in Fig. 26.

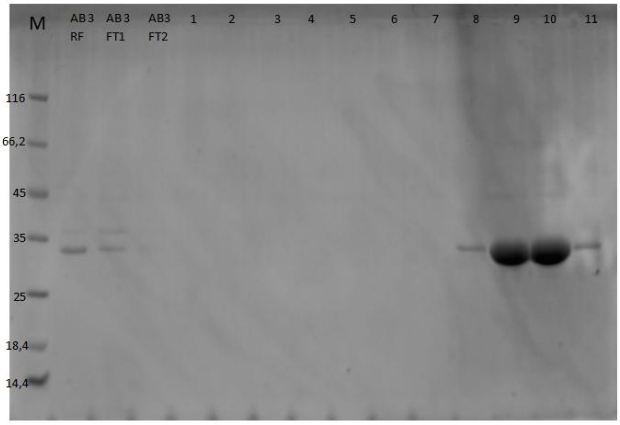


Figure 26: *SDS-PAGE of the IMAC purification. The gel shows the loaded sample containing the refolded protein (RF), the flowthrough (FT) and the collected fractions (1-11).*

The result suggests that some of the Api-Bet fusion constructs did not bind to the His-trap column (FT1) because it may not be properly folded. The eluated Api-bet 3-THX is found in the fractions 8-10 with a high concentration (2mg/mL) in the fractions 9 and 10 with less impurities. These fractions were pooled and stored at 4°C.

4.4.2.3 Cleavage

Api-Bet 3-THX contains a PreScission Protease cleavage site for splitting the Api-Bet 3 protein from the Thioredoxin part. Because the refolding of the Api-Bets without tag seemed to be much faster and less expensive because no protease cleavage is necessary, this purification of the fusion construct with the cleavage step was not followed up. Additionally, the tag-free and pre-cleaned Api-Bet 3 is still available and could easily be produced in the same way as the other three constructs.

4.4.3 Api-Bet 2 and 4 with His-tag

4.4.3.1 Refolding

The refolding of the two His-tag variants was successful via rapid dilution in 50mL buffer with 20% glycerol. A lower glycerol concentration lead to immediate precipitation of the protein. The whole batches of Api-bet 2 and 4 were taken for an IMAC purification.

4.4.3.2 IMAC

The IMAC was performed as described in chapter 3.4.3. The loaded protein samples, the flowthrough during the washing steps and all protein containing fractions, previously checked for protein concentration were taken for SDS-PAGE to monitor the purification process.

Fig. 27 shows the purification process of Api-Bet 2.

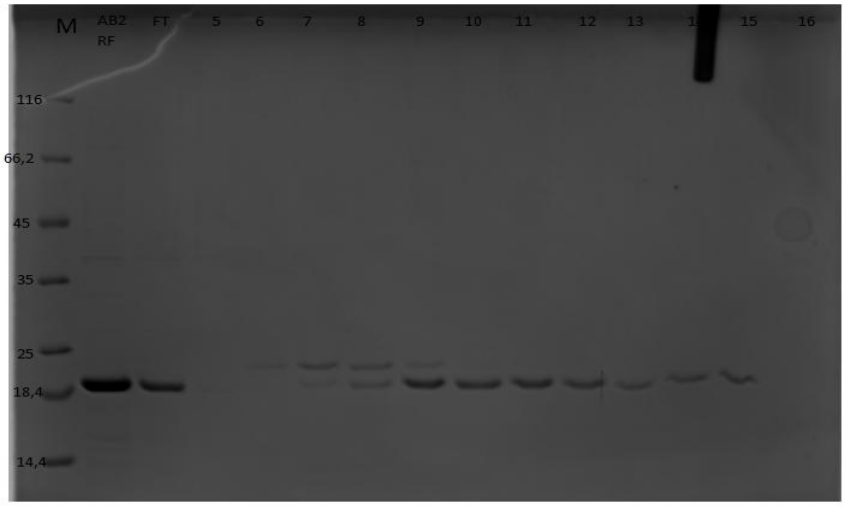


Figure 27: *SDS-PAGE of the IMAC purification. The gel shows the loaded sample containing the refolded protein (RF), the flowthrough (FT) and the collected fractions containing protein (5-16).*

The gel shows that Api-Bet 2 is found in the flowthrough in a relatively large amount. It may be that this is not a properly folded species of Api-Bet 2. The bound construct can be found in the fractions 7-15 with fewer impurities in the fractions 9-15. These fractions were pooled and stored at 4°C.

The result of the purification of Api-Bet 4 is shown in Fig. 28.

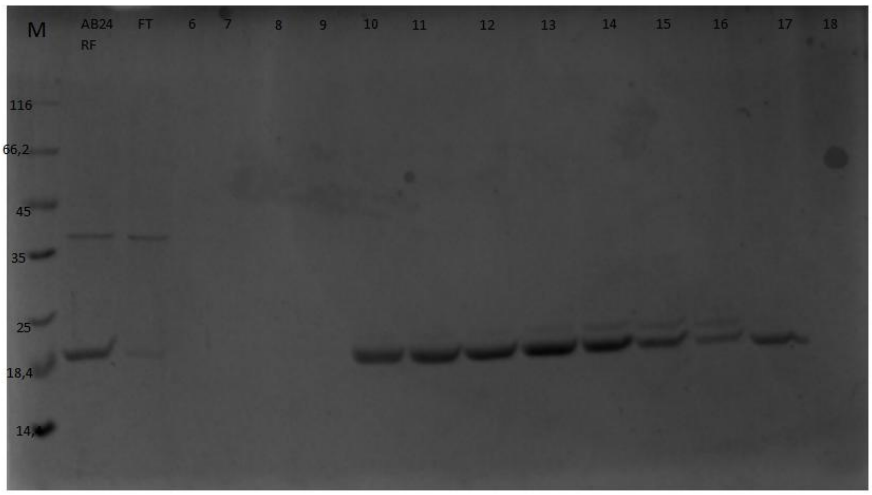


Figure 28: *SDS-PAGE of the IMAC purification. The gel shows the loaded sample containing the refolded protein (RF), the flowthrough (FT) and the collected fractions containing protein (6-18).*

The SDS-gel indicates that only a small amount of the target protein is found in the FT compared to Api-Bet 2. Api-Bet 4 can be found in the elution fractions 10-17 with fewer impurities. These fractions were again pooled and stored at 4°C.

4.4.3.3 Cleavage

These two constructs have a cleavable His-tag linked by Factor Xa-cleavage site. The refolding of the constructs and the IMAC was performed in a buffer containing 20% glycerol. In this condition, the Factor Xa protease is not active. A lower glycerol concentration in the buffer is not possible because the Api-Bet-His constructs are no more soluble. Therefore, a cleavage of the His-tag could not be performed.

Because Api-Bet 2 and 4 could be cleaned easily by refolding and a following size exclusion chromatography, the purification strategy with the His-tag was abandoned.

4.4.4 SDS-PAGE Api-Bet constructs

To compare all clean and tag-free Api-Bet constructs, a SDS-PAGE was performed. The gel (Fig. 29) shows the refolded and purified Api-Bet 1, 2 and 4 and the pre-cleaned and cleaved Api-Bet 3 from Vienna.

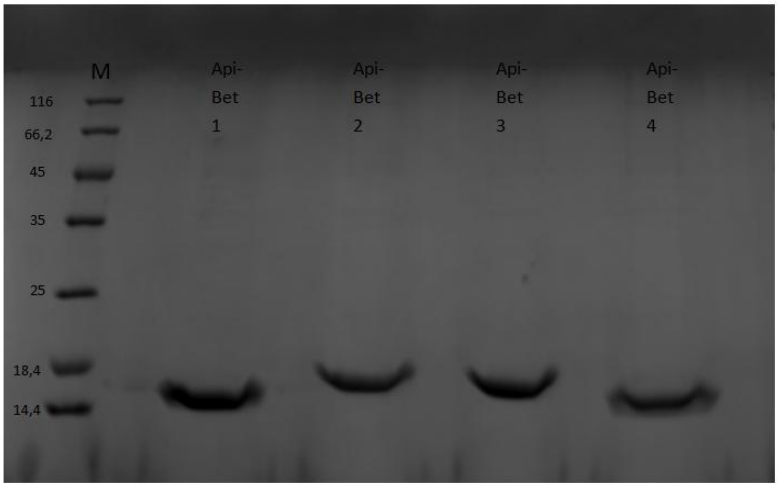


Figure 29: *SDS-PAGE of the purified Api-Bet constructs. Api-Bet 1, 2 and 4 were self-expressed, refolded and purified and Api-bet 3 is the pre-cleaned chimera from Vienna.*

The gel shows well-purified and clean protein in each case. Interestingly, the four constructs show slightly different characteristics on the SDS-PAGE.

4.5 Circular dichroism analysis

4.5.1 Full spectra measurement

All four Api-Bet constructs show a folded state in the CD spectrum (Fig. 30). The spectra of Api-Bet 1, 2 and 4 overlay nicely and have the same shape. The minima and the maximum nearly have the same values and only shift in a minimal way. Only the spectrum of Api-Bet 3, the pre-cleaned one from the cooperation group in Vienna deviates from the other three. The shape of this spectrum is similar to the other ones, but there is a shift resulting in a much lower maximum and a higher minimum value.

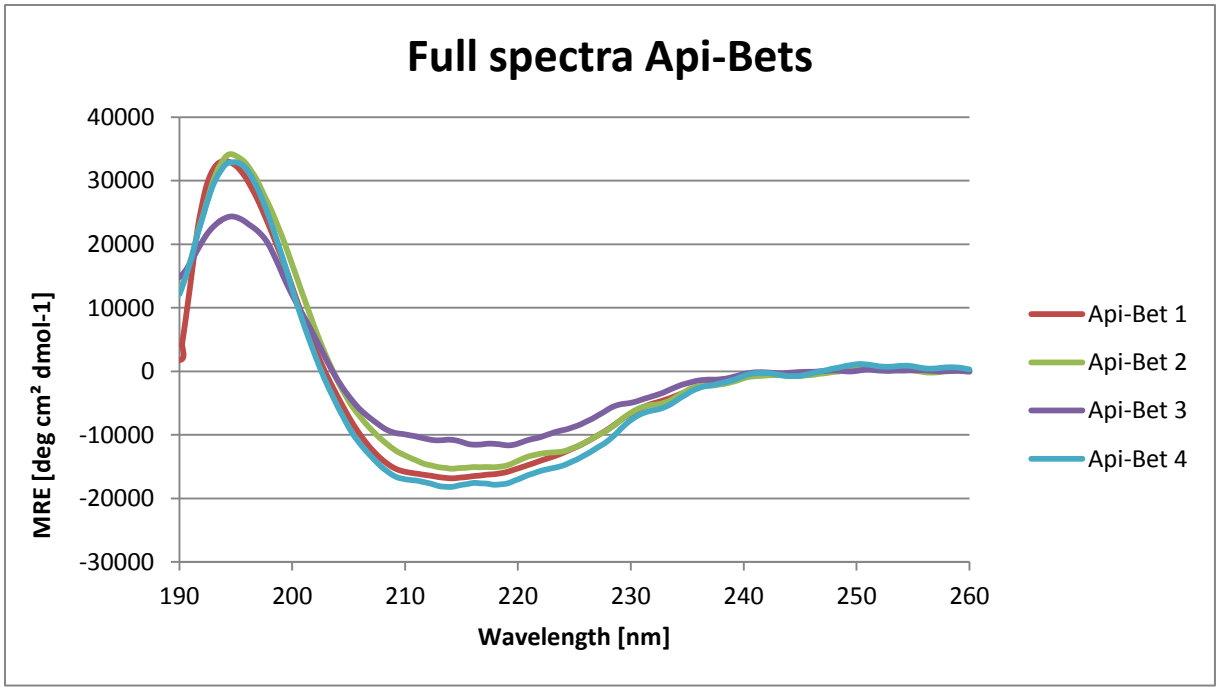


Figure 30: *Full spectra at RT of the four chimeras Api-Bet 1-4. The spectra were measured from 60-190 nm and normalized into Mean Residue Ellipticity (MRE).*

Although, the protein concentration used for normalization was checked twice, the shift in the spectrum looks like a concentration effect. It is also possible that this chimera is not completely properly folded because of freezing and thawing, which can have an effect on the folding state of a protein, despite the fact that only the supernatant after centrifugation was used for the CD-spectroscopy to remove possible precipitated protein.

Aside from the fact that the spectrum of Api-Bet 3 shifts in the values, the spectra show the expected curves like it was already published of the chimeras and the parent proteins Api g 1 and Bet v 1 (supplemental (chapter 9)) [25]

The shape of the spectra as well as the values of the molecular ellipticity indicate properly folded protein with the characteristics of a protein with a high ratio of β-sheets and also α-helical parts.

4.5.1.1 Secondary structure determination

The analysis with the dichroweb database for the secondary structure element composition reveals the following results (Table: 9).

Table 10: *Dichroweb analysis of the full spectra at RT for the four chimeras to determine the secondary structure element composition.*

Chimera	Helix1	Helix2	Strand1	Strand2	Turn	Unordered	Total
Api-Bet 1	0.37	0.11	0.14	0.10	0.10	0.17	0.99
Api-Bet 2	0.33	0.15	0.11	0.07	0.17	0.16	0.99
Api-Bet 3	0.20	0.12	0.15	0.09	0.19	0.25	1
Api-Bet 4	0.37	0.17	0.12	0.08	0.12	0.14	1

The analysis results indicate that Api-bet 1, 2 and 4 nearly show the same folding state and composition of the main secondary structure elements with a mixture of α-helical and β-sheet containing proteins as expected. The unordered part of about 16% is in a normal range for folded protein with many flexible loop regions. Api-Bet 3 shows a weaker folded state in this analysis with a higher percentage of unordered parts which also could be seen in the full spectra (Fig. 30).

4.5.2 Temperature scans

The temperature scans at 208nm were nonviable because at this wavelength no melting process could be detected. Success came by taking the measured data points at 220nm from each of the 15 full spectra at every 5° during the heating process. These data points were plotted and fitted into a sigmoid model. Theses plots are shown in Fig. 31.

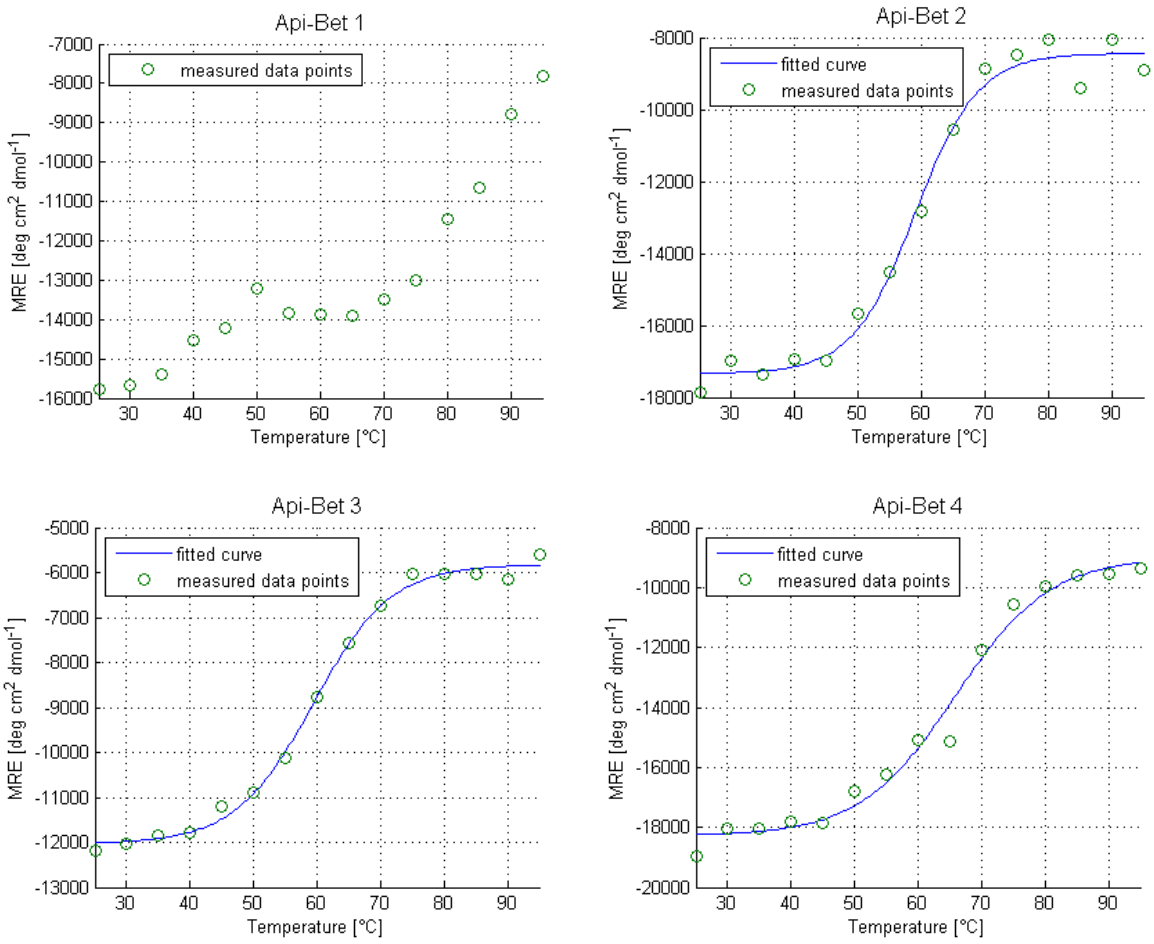


Figure 31: *Fitted melting curves based on the data points at 220nm from the full spectra from 25-95°C in 5° steps for the four Api-Bet chimeric constructs. The green points indicate the measured data points and the blue curve shows the sigmoid fit.*

Api-Bet 2, 3 and 4 show a well defined melting process that could easily be fitted into a sigmoid model. Therefore, the melting temperature could be calculated. Api-Bet 2 and 3 show a T_m of about 60°C and Api-Bet 4 is melting over a broader range with a T_m of about 65°C. The published T_m for Bet v 1 is 61°C [38] and about 70°C for Api g 1 [39] also determined by CD-spectroscopy.

Api-Bet 1 deviates from the other chimeras by showing a different melting behavior. It shows a change of the fold between 25-55°C and between 70-95°C. Between 50°C and 70°C, the signal is nearly constant. This indicates that Api-bet 1 has two melting points during the heating process. Because the H(T) signal of the CD-detector did not show any abnormal changes during the measurement compared to the other performed scans, a measuring error seems rather unlikely, but the scan should be repeated.

To monitor the melting process, the full spectra at 25°C, 55°C and 95°C are plotted in Fig. 32 for each construct.

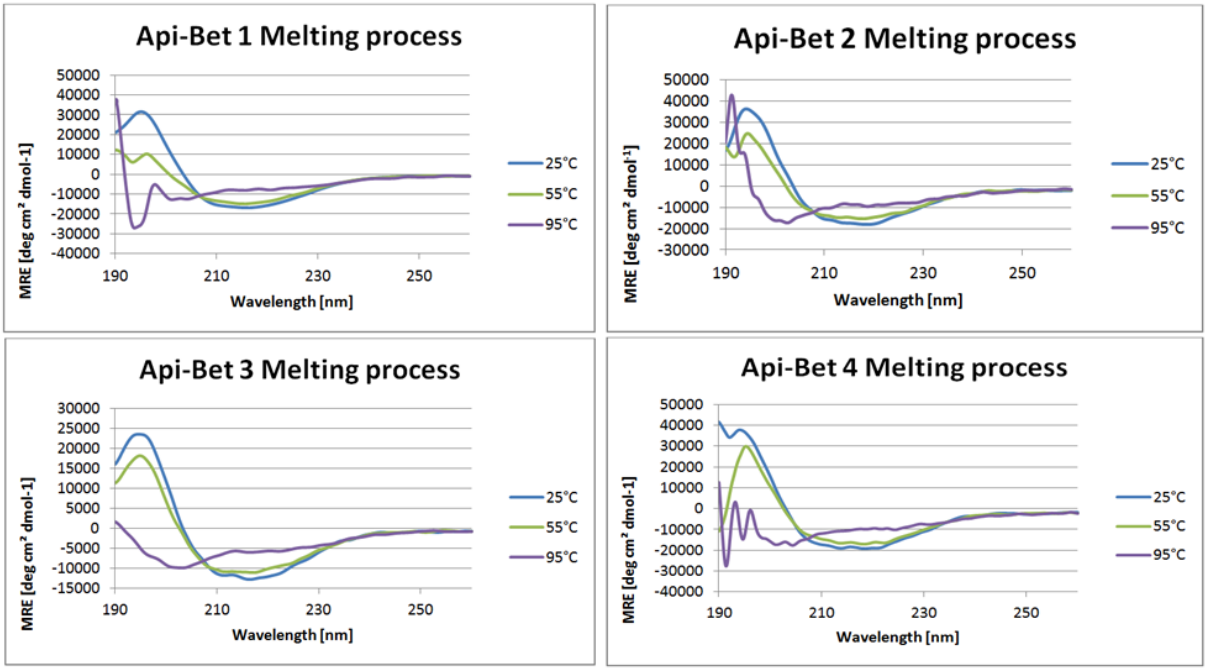


Figure 32: *Melting process of the four chimeras by overlaying the full spectra at 25°C (blue), 55°C (green) and 95°C (purple)*

The overlay of the full spectra at 25°C and 55°C show a slight shift that indicates that already some small parts of the structural integrity were lost, but the proteins are still mainly folded. The spectra 95°C indicate that all Api-Bet constructs are completely defolded at this temperature. The complete melting process of Api-Bet 1 by overlaying all 15 full spectra is plotted in the supplemental material (chapter 9)

4.5.3 Refolding capacity

Figure 33 shows the refolding capacity of each chimera by overlaying the full spectra at RT before (blue curve) and after (green curve) heating to 95°C. As shown before, at 95°C, the majority of the protein is completely unfolded and the data shows that a huge part of the proteins show a folded sate again after cooling down to room temperature.

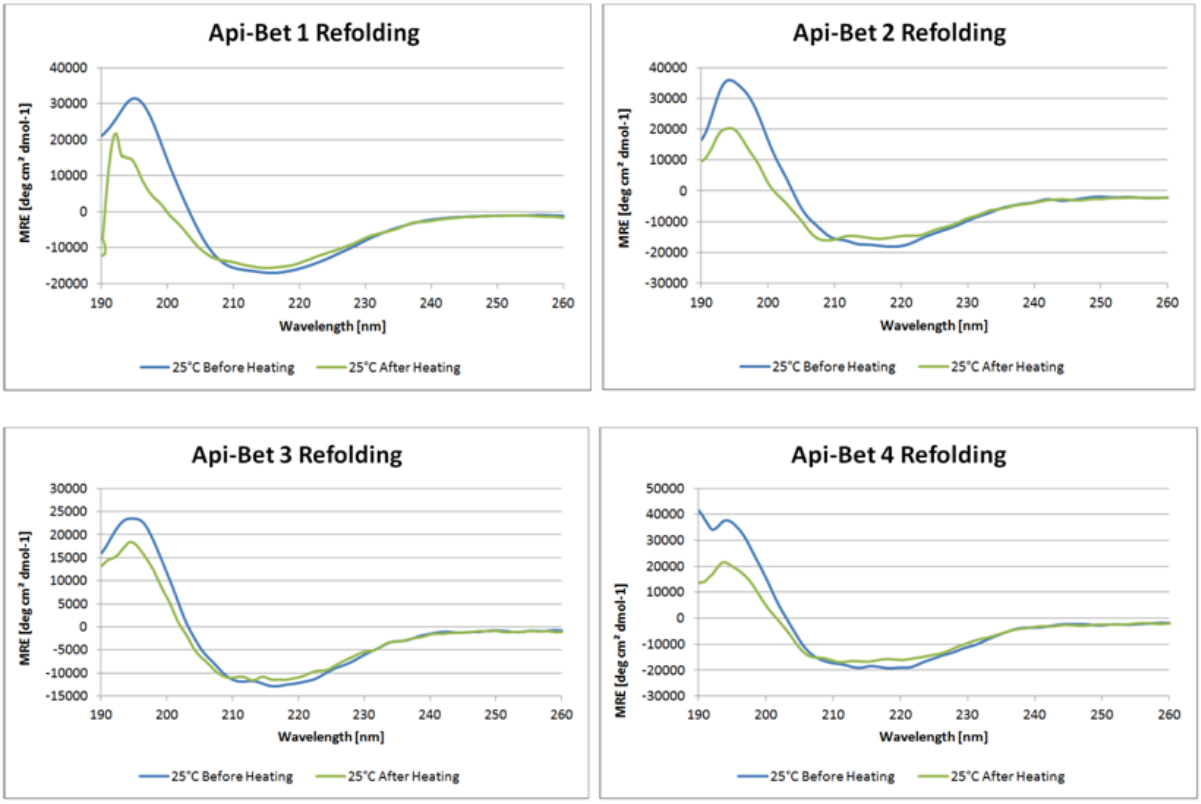


Figure 33: *Overlay of the full spectra at RT before (blue) and after (green) heating for each construct indicating the refolding capacity.*

The overlays of the spectra indicate that the constructs do not refold completely, because of the shifts in the spectra, especially the lower maximum. It may be possible that a small part of the protein sample precipitated during the heating process and did not refold, instead it stays completely unfolded. The rest of the protein, the majority of the sample, refolded completely and therefore the shift in the spectra is maybe caused by a loss of concentration. The deviations between 190-195 nm can result from detector noise caused by the salts of PBS-buffer.

4.6 Crystallization

The only hit in the crystallization screen setups were found in the PEG/ION screen in condition 32 for Api-Bet 1. This condition yielded one crystal in a mainly precipitated drop.

Further optimization lead to plenty needle-shaped crystals as shown in Fig. 32.

Figure 34: *Protein crystals of Api-Bet 1 in the optimized PEG/ION condition 32*

All crystals were checked for their diffraction on the home source but none of the optimized crystals diffracted more than 3 Å and therefore better than the main crystal found in the screen setup with a diffraction of 2.8 Å (Fig. 35). So this main crystal was sent for a full data set collection to the ESRF. The diffraction at the synchrotron was 2.65 Å.

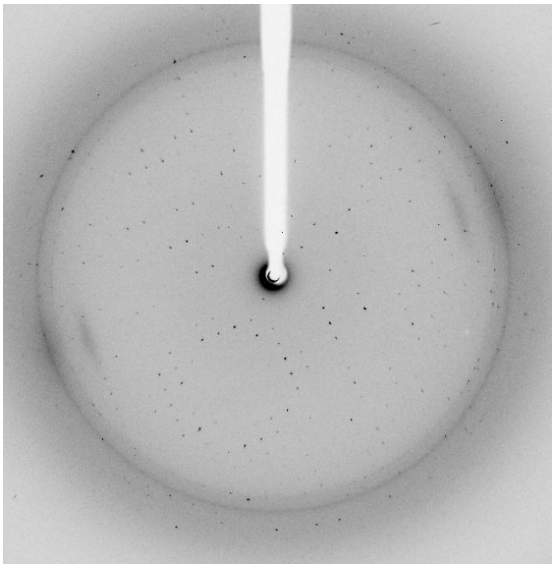


Figure 35: *Api-Bet 1 crystal diffraction at the home source.*

4.7 X-Ray analysis

The first crystal gave a full dataset at ESRF. The crystal was found to be

$P4_122$ with a cell of 52.6 52.6 131.7 90 90 90

The structure elucidation was elusive due to problems with molecular replacement.

4.8 Molecular Replacement

For solving the phase problem, molecular replacement was the method of choice because of the high similarity of Api-Bet 1 with its parent proteins. But neither Api g 1 nor Bet v 1 gave a refineable solution. To get rid of the problem, Api g 1 was split into the main secondary structure parts, the

main α-helix (C-Terminus) and the β-sheet part. The flexible loop regions were deleted and the two structure elements were taken separately for molecular replacement (Fig. 36).

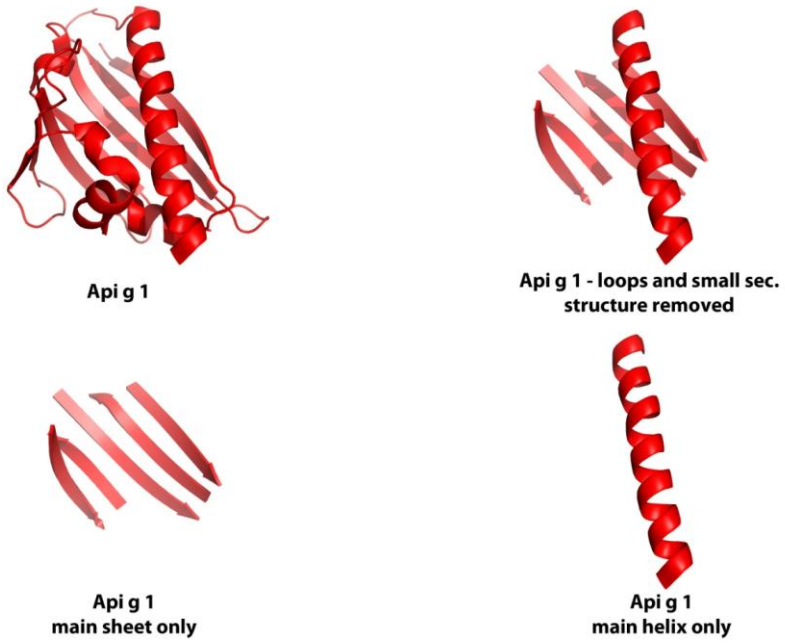


Figure 36: *Splitting of the secondary structure elements of Api g 1 used for molecular replacement*

This procedure was successful and gave a solution that could be used for the refinement and the solving of the structure of Api-Bet 1.

4.9 Structural Refinement

The building and refinement of the structure was started with a Poly-Ala model of the two major secondary structure elements that were used for molecular replacement. The first step was to find a starting point from which the amino acid sequence could be assigned. Several cycles of refinement lead to a stepwise elongation of the sequence in the electron density map. In some areas the protein sequence in the electron density was well defined in relation to the resolution (Fig. 37).

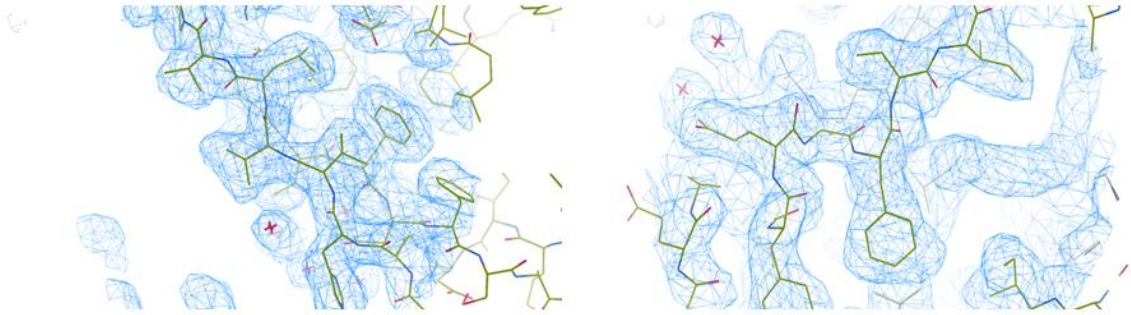


Figure 37: *Feature enhanced map (FEM) in coot at 1.5 rmsd showing the well defined areas of the Api-Bet 1 crystal structure.*

The building of the structure caused many problems during the refinement process. There was a lot of undefined and unclear electron density found that could not be filled. Additionally, some loop regions could not be built because of missing electron density leading to free open ends as shown in Fig. 38.

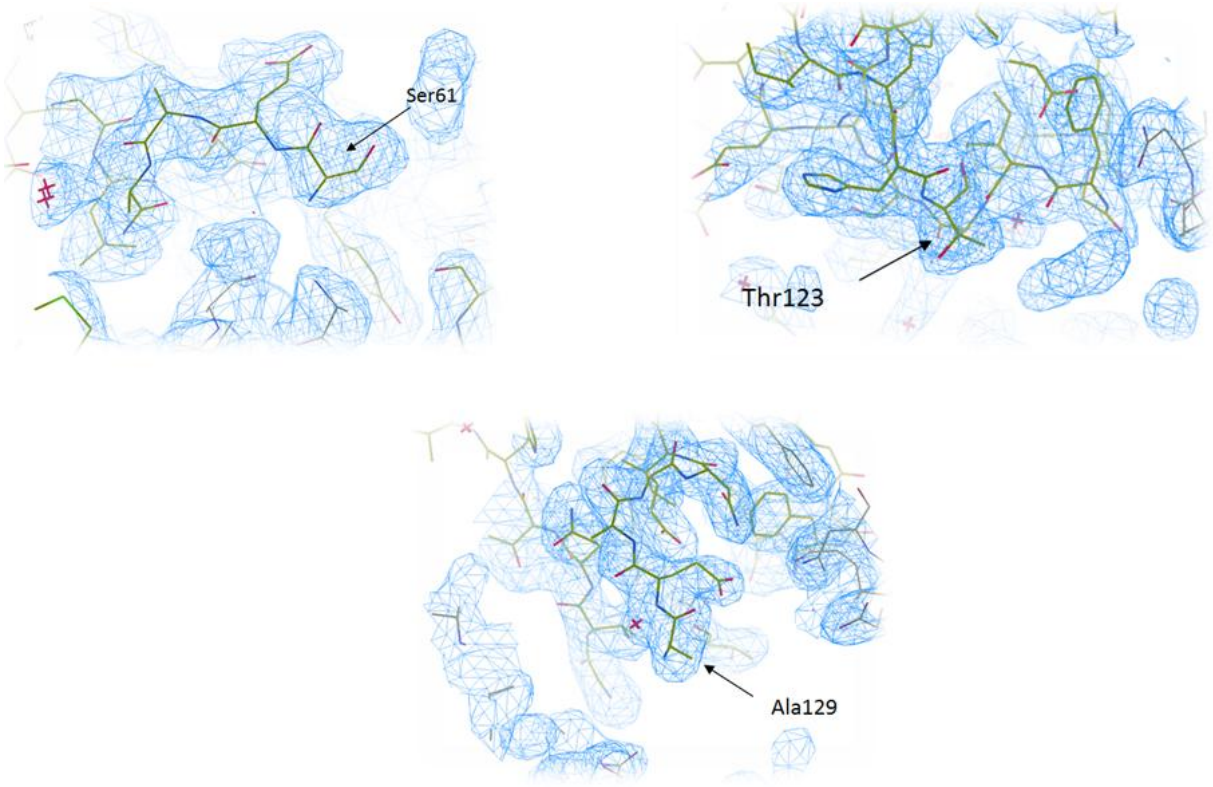


Figure 38: *Electron density map (FEM) at 1.5 rmsd showing the free open endings at Ser61, Thr123 and Ala129.*

This lead to the fact that the flexible loop between residue 123 and 129 could not be built. Additionally, the sequence between residue 30 and 63 could not be assigned, except for a small sequence part in between, because the corresponding electron density was not visible in the map. Therefore, only 124 of the 154 amino acid residues of the Api-Bet protein could be seen in this crystal structure. Although, a lot of undefined electron density is still left without an obvious possibility to fill it.

The structure was refined to an R_{free} of 0.35. This value is quite high and could possibly be reduced further by filling the empty electron density areas. The final structure and refinement statistics calculated with the final pdb-file, the mtz-file and the unmerged data file are shown in table 11.

Table 11: *Data collection, merging and refinement statistics of Api-Bet 1*

Data collection and Refinement Statistics	
Resolution range (Å)	48.78 - 2.649 (2.746 - 2.651)
Space group	P 4_1 2 2
Unit cell	52.52 52.52 131.69 90 90 90
Total reflections	99118 (9045)
Unique reflections	5259 (566)
Multiplicity	17.0 (16.0)
Completeness (%)	90.04 (91.99)
Mean I/sigma(I)	25.22 (3.87)
Wilson B-factor	48.80
R-merge	0.8552 (1.476)
R-meas	0.8828
CC1/2	0.719 (0.783)
CC*	0.915 (0.937)
R-work	0.2759 (0.4045)
R-free	0.3519 (0.4204)
Number of non-hydrogen atoms	959
macromolecules	926
ligands	0
water	33
Protein residues	124
RMS(bonds)	0.012
RMS(angles)	1.63
Ramachandran favored (%)	75
Ramachandran outliers (%)	11
Clashscore	41.76
macromolecules	
ligands	0
solvent	44.40

4.10 Structure

The so far solved structure of Api-Bet 1 found in the crystal is shown in Fig. 39. It shows many unexpected characteristics that are shown in the Figure in different views. The graphics were designed with PyMOL.

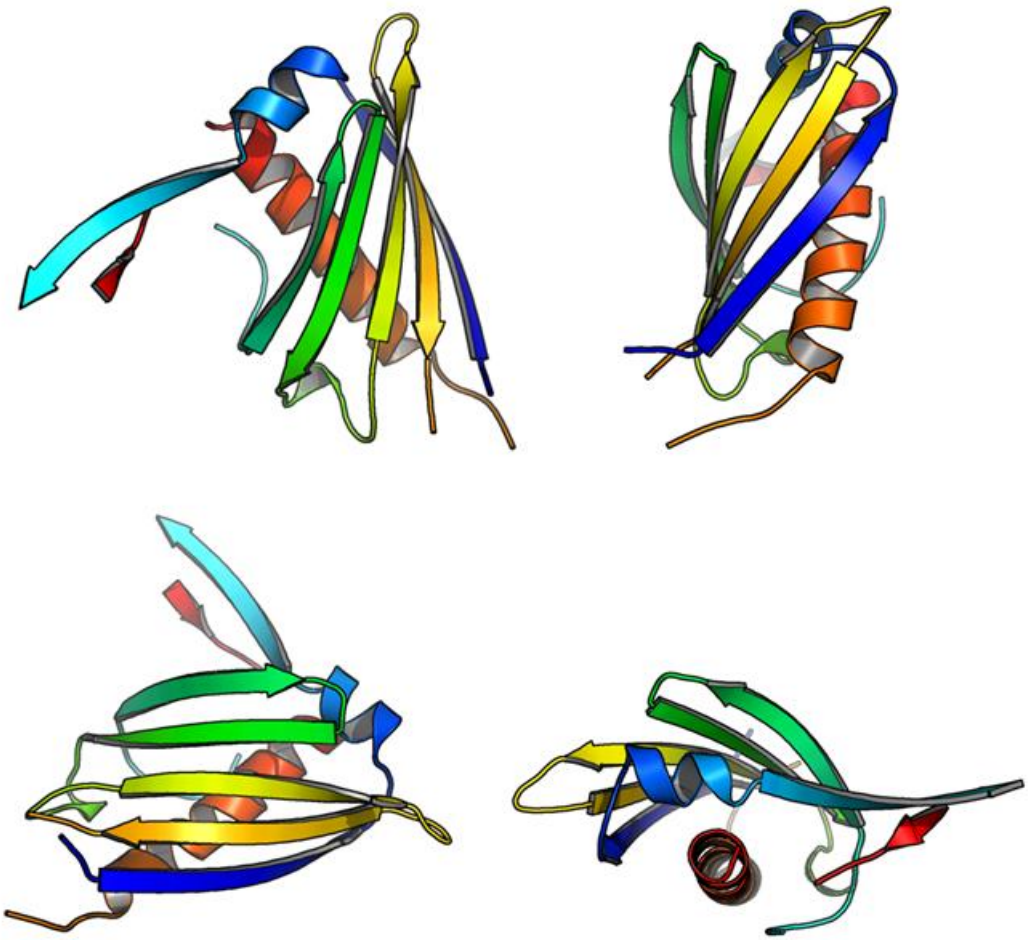


Figure 39: *Crystal structure of Api-Bet found in different views.*

Only five of the seven antiparallel β-sheets around the hydrophobic cavity are formed. Instead, two β-strands seemed to grow outside the entire structure.

The investigation of the symmetry mate revealed the following result (Fig. 40):

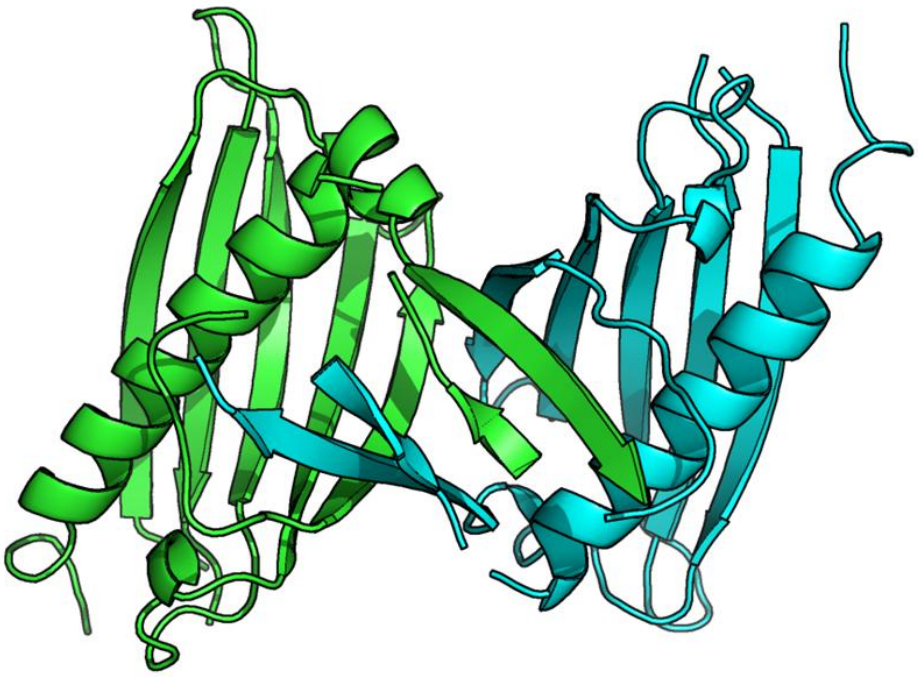


Figure 40: *Dimeric arrangement of Api-Bet 1 in the crystal structure.*

In the crystal structure, Api-Bet 1 forms a dimer with the symmetry molecule with the β-sheet parts growing into the neighboring molecule and vice versa.

For a better illustration, the dimer found in the crystal is compared with the parent structure of Api g 1 with the mutations of Api-Bet 1 shown in blue (Fig. 41).

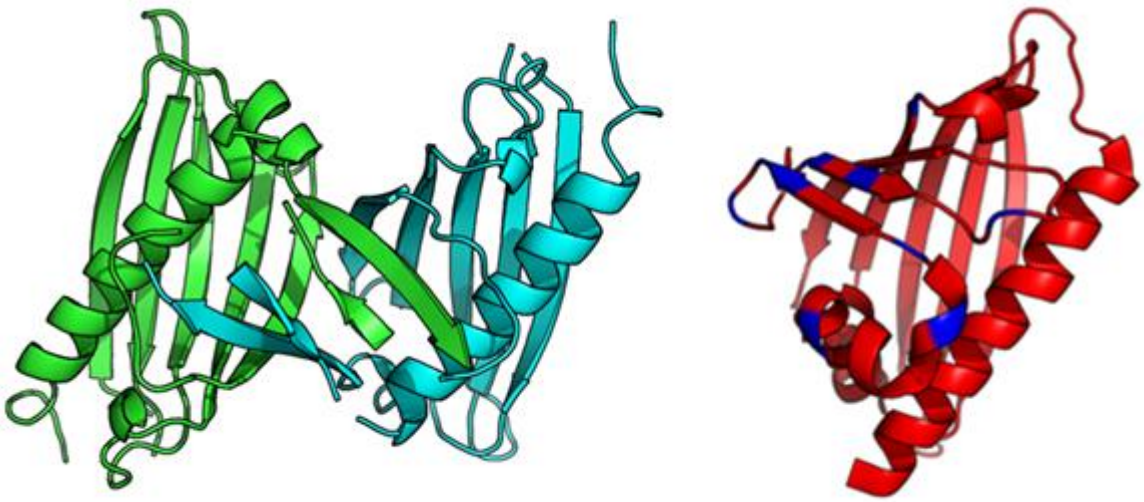


Figure 41: *Dimeric arrangement of Api-Bet 1 in the crystal (green and cyan) and the parent structure of Api g 1 (red) with the assigned mutated areas of Api-Bet 1 (blue).*

The comparison of these structures revealed that in Api-Bet 1 the second small α-helix (α2) is not formed. Instead, the sequence in this part forms a β-sheet connecting the symmetry molecule. This phenomenon of switching of a second structure element into another is called domain swappimg. Another different aspect found in comparison is that the large main α-helix (α3) moved much closer towards the β-strands leading to a collapse of the hydrophobic cavity. Interestingly, the deviations occur in the mutated areas of the small helices, the P-loop and the 2. and 3. β-sheet that are almost missing completely in their native conformation.

Nevertheless, the domain swapped area is the most impressive part of the structure. Although there are many undefined areas, the domain swapped arrangement with the symmetry mate can clearly be seen in the electron density map (Fig.42).

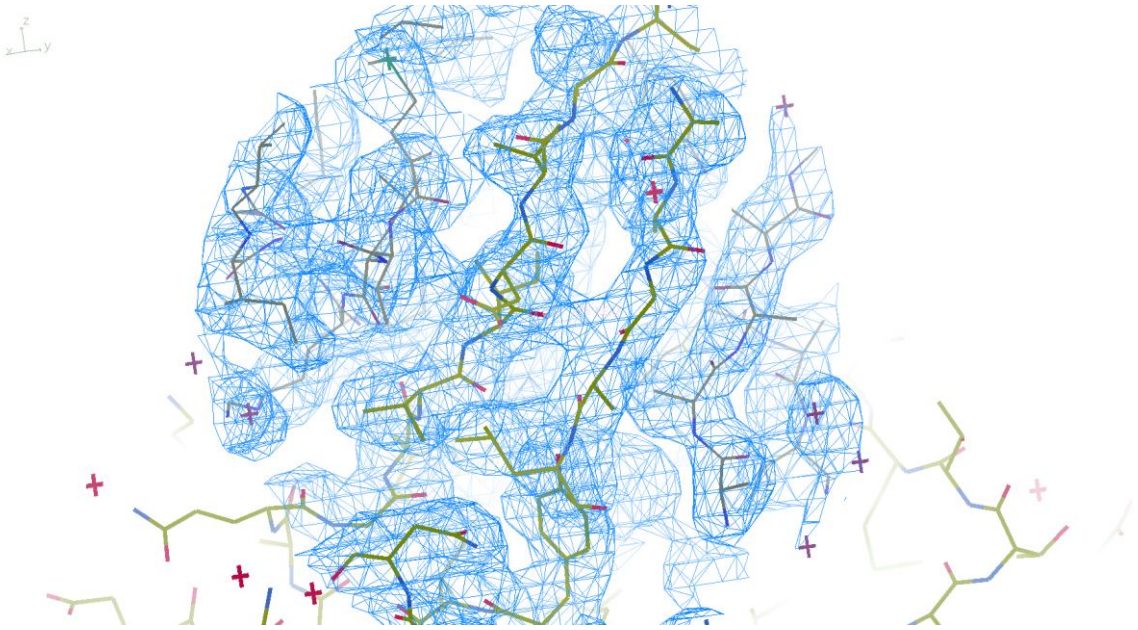


Figure 42: *Electron density map (FEM) at 1.5 rmsd of the domain swapped area of Api-Bet 1 with the symmetry molecule in the crystal.*

A detailed comparison with Api g 1 by a direct alignment of both structures revealed that the five defined β-strands (β1 and β4-β7) are overlaying very well (Fig. 43, left view).

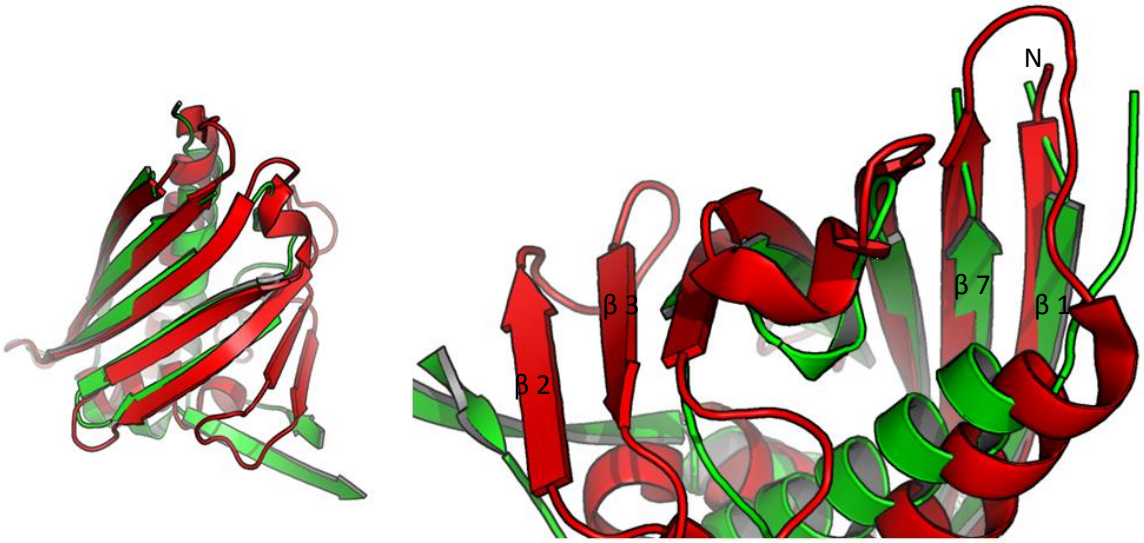


***Figure 43**: Comparison of Api-Bet 1 in the crystal and Api g 1 by direct alignment of both structures. The left view shows the β-stranded areas and the right view shows the main differences of both structures.*

The right view in Fig. 43 shows the main differences of both aligned structures. The two main helices are shifting and the β-sheets 2 and 3 with the connecting loops are not found in the crystal structure of Api-Bet 1. This part is the major area of the domain swap. Due to refinement problems, the loop between β7 and α3 could not be built.

Fig. 44 shows the detailed comparison of the domain swapped region.

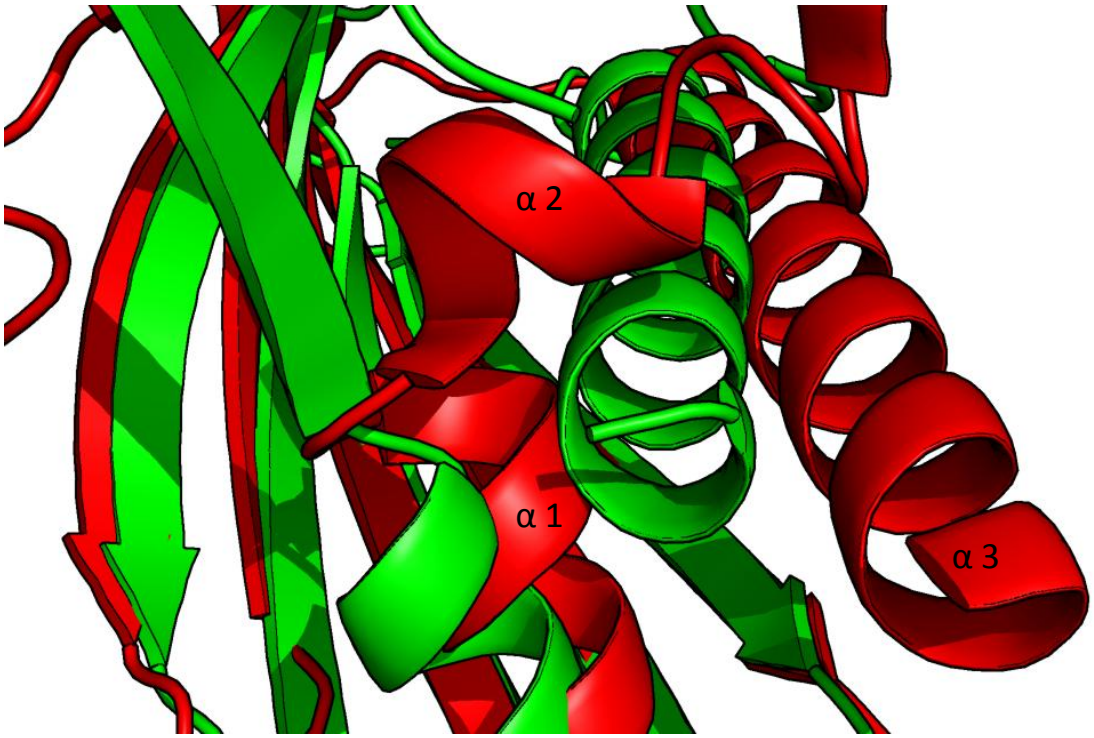


***Figure 44:** Detailed comparison of Api-Bet 1 and Api g 1 in the domain swapped region by direct alignment.*

The alignment shows that the first small helix after β1 is still well defined but instead of forming the second small helix subsequently, the sequence of Api-Bet 1 forms a β-sheet connecting to the neighboring molecule. Again, the clash of the major α-helix near to hydrophobic cavity is presented resulting in a decrease in the size of the hydrophobic cleft.

Additionally, the Api-Bet 1 structure is aligned with calculated Phyre²-model of Api-Bet 1 (Fig. 45)

Figure 45: *Comparison of the Api-Bet 1 structure (rainbow) and the calculated Phyre²-model of Api-Bet 1 (magenta)*

The alignment with the Phyre²-model shows nearly the same deviations as the alignment with Api g 1. Again the domain-swapped region and the shift of the major helix are predominant in the alignment.

The surface illustration of the Api-Bet 1 dimer in comparison with the surface of Api g 1 is shown in Fig 46.

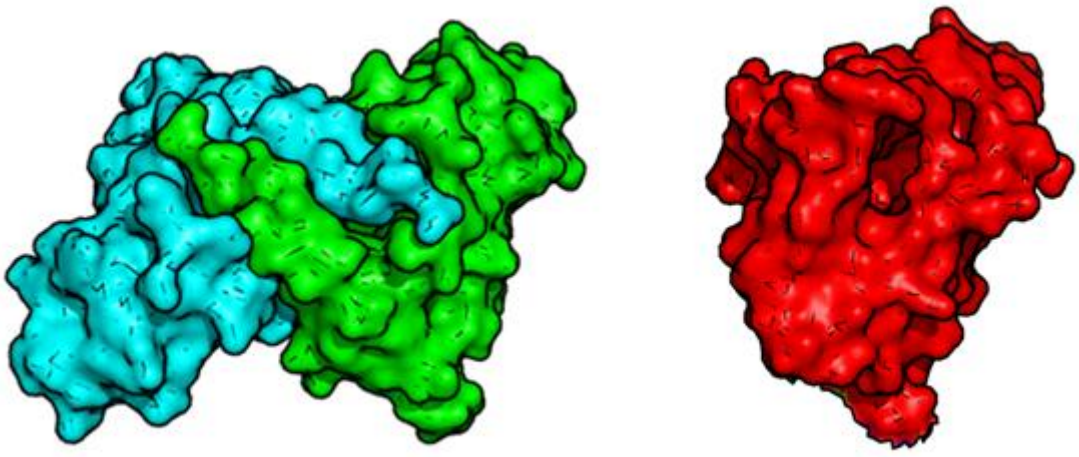


Figure *46: Surface illustration of the Api-Bet 1 dimer in the crystal structure (green and cyan) and of Api g 1 (red)*

The dimer of Api-Bet 1 shows a very compact arrangement. The comparison with the surface of Api g 1 reveals that the β-sheet back is very similar.

To evaluate the structure regarding the correctnes, a b-factor was calculated shown in Fig. 47.

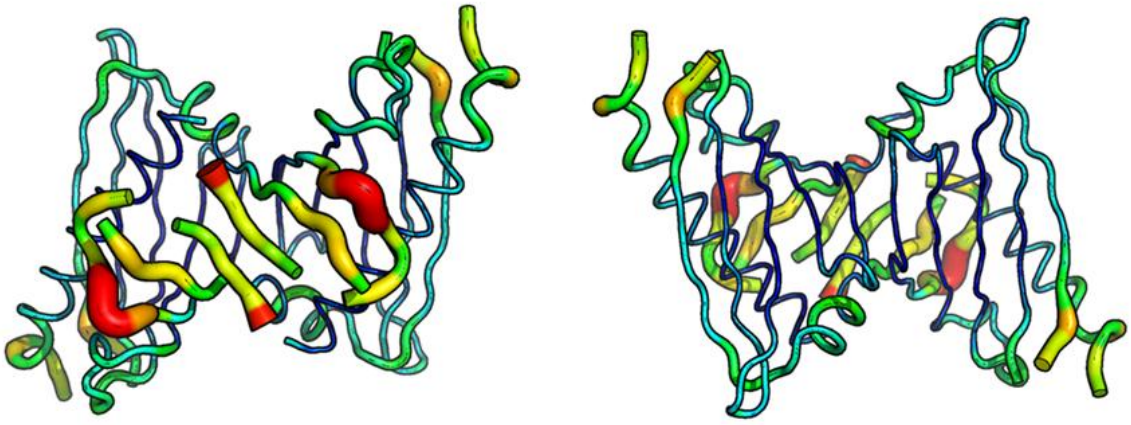


Figure 47: *B-factor model of the Api-Bet 1 dimer.*

This model shows well defined areas with low b-factors in lean tubes with a coloration of blue to green and high b-factor regions that have a lower validity in broader tubes in yellow to red. Here it can be seen that the major α-helix and the β-sheets (β1 and β4-β7) are well defined with low b-factors. The most undefined region in this structure is the domain swap with the junction of the small helix to the domain swapped sheet.

Because of the presence of the protein in a dimeric arrangement in the crystal, a size exclusion chromatography of the pre-cleaned protein sample was performed with a Superdex 200 column (GE Healthcare) with a flow of 0.5 mL/min in a Na-citrate buffered 100mM NaCl solution (pH 6,4) to obtain the size of the protein and its aggregation state in solution (R_{hyd}). A standard run was also performed in the same setup afterwards.

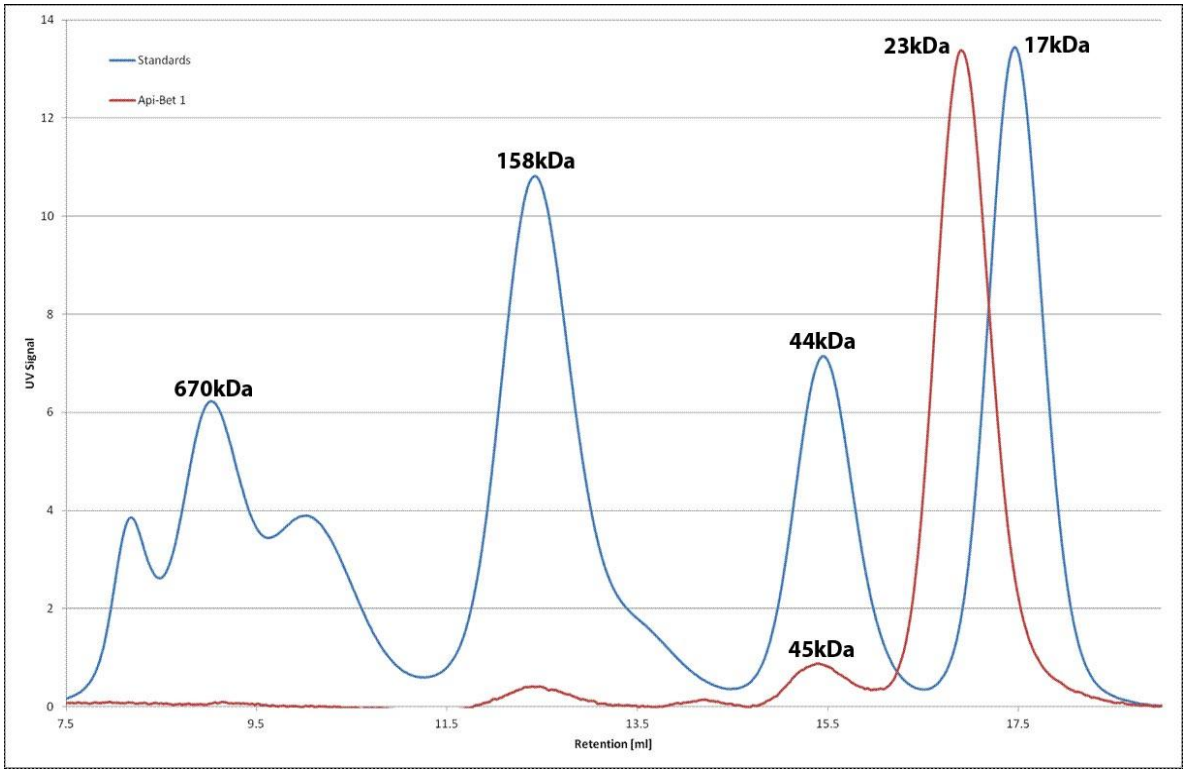


Figure 48: *Size exclusion chromatogram of Api-Bet 1 (red) and standards (blue).*

Interestingly, the chromatogram of Api-Bet 1 shows, beside a main peak at 23 kDa, a second weak peak fraction at 45 kDa. The peak at 23 kDa corresponds to monomeric protein taking into account that Api-Bet has a larger hydrodynamic radius than expected for a 17 kDa protein. The second peak at 45 kDa can represent a dimer fraction, which means that also in solution a small amount of the sample is present in a dimeric state.

4.11 SAXS measurement

The most important aspect of this experiment was the determination of the oligomerization state of Api-Bet 1 because of the so far illustrated dimeric arrangement in the crystal structure. The SAXS measurement was necessary to see if the newly expressed and purified Api-Bet 1 also shows this effect in solution at high concentration.

After the initial buffer substraction, the SAXS data was merged, evlauated and adjusted with the samples at the different concentrations to reveal a well-fitted diffraction curve. To evaluate the oligomerization sate and the overall shape of the protein particles in solution, a theoretical diffraction curve of the monomeric Phyre²-model of Api-Bet 1 was calculated and overlayed with the measured curve (Fig.41) and the fit of these two curves was calculated.

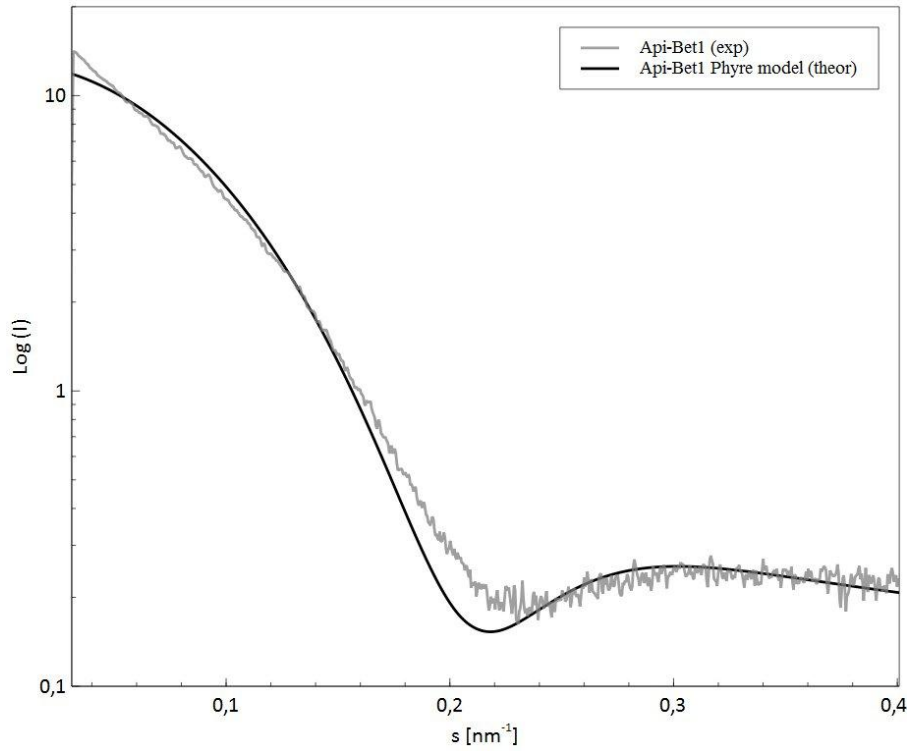


Figure 49: *Measured diffraction curve of Api-bet 1 (grey) and theoretically calculated diffraction curve of the Api-Bet 1 Phyre model (black)*

Both curves, the measured and the theoretical one, are overlaying very well. Additionally, the Chi-value of 3.503 indicates that both models fit well, which means that the measured protein particles are similar to the Phyre²-model of Api-Bet 1.

Beside the Phyre²-model, a theoretical diffraction curve of the crystal structure of Api-Bet 1 (chapter 4.10) in the monomeric state was calculated and again overlayed with the measured diffraction curve of Api-Bet 1 shown in Fig. 42.

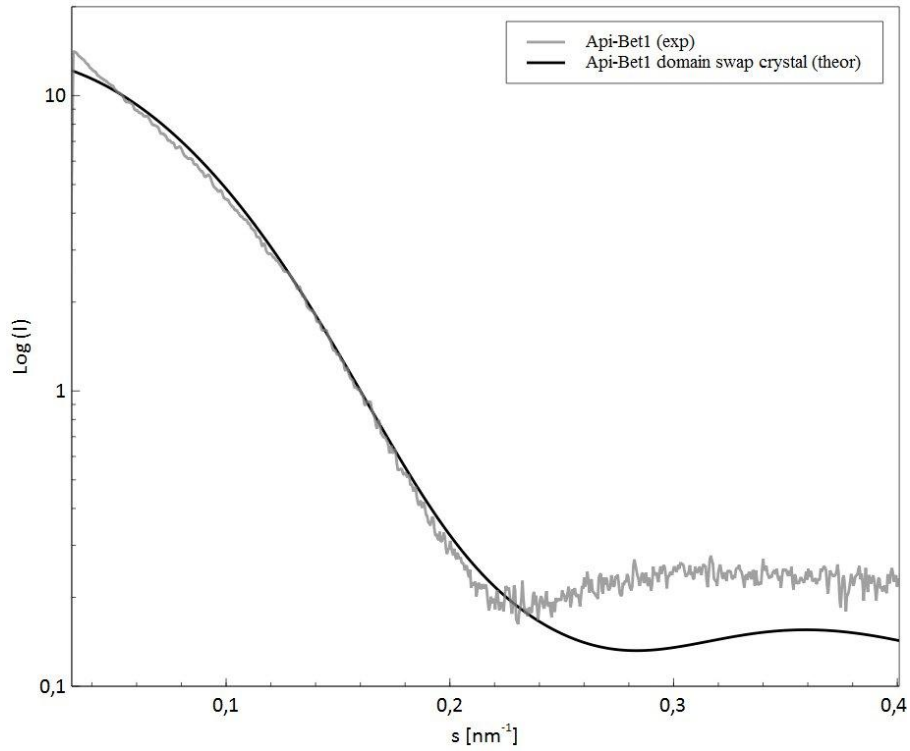


Figure 50: *Measured diffraction curve of Api-bet 1 (grey) and theoretically calculated diffraction curve of the Api-Bet 1 crystal structure in the monomeric form (black).*

It is obvious that both curves are deviating from each other. The Chi-value of 4.184 shows that the fit of the two curves is not very well compared to the first one of the Phyre²-model. This indicates that the Api-Bet 1 seems to be properly folded and monomer and that the sample does not show the characteristics of the crystal structure in solution.

Another process to evaluate the measured protein particles regarding their size is the calculation of their measured molecular weight in solution. For this, a reference protein is needed. In this case, BSA as a standard protein was measured in the same setup. With the MW of BSA, the measured I(0) values of BSA and the sample (Api-Bet 1), the measured MW of the Api-Bet 1 particles can be calculated with the following equation:

$$\boldsymbol{MW_{Api}} = \frac{\mathbf{MW_{BSA}}}{\boldsymbol{I[0]_{BSA}}} \cdot \boldsymbol{I[0]_{Api}}$$

The insertion of the constants and measured values in the equation reveals the following solution for the theoretical molecular weight for the measured protein particle:

$$\boldsymbol{MW_{Api}} = \frac{\mathbf{66\ kDa}}{\mathbf{69.528}} \cdot \mathbf{18.035} = \mathbf{17.12}\ \boldsymbol{kDa}$$

Compared with the real molecular weight (ProtParam) of Api-Bet 1 with 16.4 kDa, the measured value of the theoretical MW with 17.12 fits very well to a monomeric protein in solution.

The SAXS measurement and the processing of the data revealed that the newly expressed and purified Api-Bet 1 is monomeric and not aggregated or dimerized in solution even at a relatively high concentration.

5 Discussion

The main aim of this work was to express and purify the four Api-Bet constructs for a further biophysical and structural characterization. All constructs could nicely be expressed in E. *coli* but previous tries with the purification of the tag-free proteins resulted in huge problems regarding the output of sufficient amounts of clean protein. Therefore, the insoluble fraction of the expression was taken for an IB purification followed by refolding of the constructs. The refolded protein was finally purified by SEC yielding soluble and monomeric protein in large amounts with less effort. Because the purification process by refolding is so much less time- and resource consuming and also yielded much more amounts of clean protein, the cloning, expression and purification of the other fusion constructs was dropped.

The folding state of the purified proteins was checked by CD-spectroscopy. All expressed and refolded chimeras (Api-Bet 1, 2 and 4) show proper folding state of a mixed α-helical and β-stranded protein as it is published also for the parent proteins Api g 1 and Bet v 1 (Supplemental). Only the pre-cleaned Api-Bet 3 sample from Vienna deviates in the spectrum which indicates a less proper folding which may be caused by freezing , thawing and long-term storage.

So far the crystal structure of Api-Bet 1 could be solved but with many problems regarding the processing and the refinement of the structure data. The MR with the structure data of the parent proteins was not possible. Because in the crystal structure the hydrophobic cavity collapsed by a clash of the major α-helix near to the β-strands, the structure shows a different architecture. This is the reason why only replacing of the two major secondary structure parts in a separate way revealed success for a molecular replacement.

The structure of Api-Bet 1 could be refined to an R_{free} of 0.35. This value is still very high because of many electron density parts that could not be filled. Additionally, the sequence of 30 residues could not be built because of missing density. Therefore, the refinement statistics reveals many bad parameters. So in sum, the structure cannot be seen as completely solved.

In the crystal, Api-Bet 1 shows an unexpected domain-swapped dimeric arrangement. Interestingly, the collapsed regions are found in the parts of the mutated areas of the small helices, the β-sheets 2 and 3 and the p-loop. It seems that the introduced mutations have a destabilizing effect on the fold of Api-Bet 1. This hypothesis should be proofed with further investigations.

The b-factor model of the crystal structure of Api-Bet 1 indicates that the domain swapped region in the structure seems to be the most undefined part. But nevertheless, there seems to be no other possibility to build this region as shown by the electron density map.

The SEC-chromatogram of the batch that was used for crystallization and structure determination shows a small peak at the size of a dimer beside the main peak of monomeric Api-Bet 1. So it may be that only this small dimer species did crystallize as a dynamic process whereas the monomeric fraction precipitated. The next crystallization setups should be performed only with monomeric protein sample to see if this dimer is a side effect of crystallization.

Because of this collapsing dimerization of Api-Bet 1 in the crystal structure, the focus in the purification and characterization of newly expressed batches was the preferably accurate determination of the oligomerization state of the sample. Therefore, an Api-Bet 1 sample was sent for a SAXS-measurement. The experiment revealed that this chimera seems to be monomeric and not aggregated even at high concentrations in solution at RT. The comparison of the measured diffraction curve with the theoretical curves of the Phyre²-model and the collapsed structure indicates that the Api-Bet 1 protein particles show an expected size and shape like the native fold.

In the context of the dimeric domain-swapped arrangement of Api-Bet 1 in the crystal structure, the temperature CD-scan (chapter 4.5.3.) becomes very interesting. The melting curve shows two independent melting processes, one at about 45°C and another at about 72°C. It can be that the mutations revealed a destabilization by shuffling of structural parts even at relatively low temperatures. Maybe this is the critical point where the protein starts interacting with a neighbor protein and dimerizes by forming the domain swap that is enabled by the mutations. This can lead to a stabilization of proteins by dimerization with each other that causes a second, complete melting at high temperatures compared to the normal T_m of the other chimeras.

So in sum, much more investigation is needed to explain the recent results.

6 Outlook

The first step in the further research is to reproduce the crystal structure of Api-Bet 1. Since the chimeric constructs could be expressed and purified easily in large amounts now, a lot of crystallization setups with different protein concentrations should be tried starting with the condition that lead to a hit before. Although the generation of a new crystal structure is needed to see if the domain swap is reproducible, the aim is to solve the structures of all four api-Bet chimeric constructs.

The size exclusion chromatographies should be repeated after a certain time of storage to see if the dimerization is maybe a time dependent process.

Additionally, the CD-temperature scan of Api-Bet 1 has to be reproduced to be sure, that the abnormal melting behavior is not the result of a measuring error.

Another aspect is to search for a ligand that binds to the Api-Bet constructs. Because the parent proteins Bet v 1 and Api g 1 show differences in their ligand binding because of different architectures of their hydrophobic cavities, the binding capacities of the chimeras seems to be very interesting. Additionally, a bound ligand could stabilize the hydrophobic cleft and the connecting helices and could therefore have a protective effect regarding a possible domain swap and dimerization. Maybe in this way the monomeric protein could be stabilized.

In the case that Api-Bet furthermore appears abnormal in its biophysical behavior, computational investigation by molecular dynamics should be done to study possible influences of the introduced mutations and revealed clashes.

7 References

1. Beasley, R., et al., *Prevalence and etiology of asthma.* J Allergy Clin Immunol, 2000. **105**(2 Pt 2): p. S466-72.
2. Majkowska-Wojciechowska, B., et al., *Prevalence of allergy, patterns of allergic sensitization and allergy risk factors in rural and urban children.* Allergy, 2007. **62**(9): p. 1044-50.
3. Ring, J., et al., *Why are allergies increasing?* Curr Opin Immunol, 2001. **13**(6): p. 701-8.
4. Uzzaman, A. and S.H. Cho, *Chapter 28: Classification of hypersensitivity reactions.* Allergy Asthma Proc, 2012. **33 Suppl 1**: p. S96-9.
5. Valenta, R., *The future of antigen-specific immunotherapy of allergy.* Nat Rev Immunol, 2002. **2**(6): p. 446-53.
6. He, S.H., et al., *Mast cells and basophils are essential for allergies: mechanisms of allergic inflammation and a proposed procedure for diagnosis.* Acta Pharmacol Sin, 2013. **34**(10): p. 1270-83.
7. Poulsen, L.K. and L. Hummelshoj, *Triggers of IgE class switching and allergy development.* Ann Med, 2007. **39**(6): p. 440-56.
8. Jones, M., *Understanding of the molecular mechanisms of allergy.* Methods Mol Med, 2008. **138**: p. 1-15.
9. Pope, S.M., et al., *IL-13 induces eosinophil recruitment into the lung by an IL-5- and eotaxin-dependent mechanism.* J Allergy Clin Immunol, 2001. **108**(4): p. 594-601.
10. Winter, W.E., N.S. Hardt, and S. Fuhrman, *Immunoglobulin E: importance in parasitic infections and hypersensitivity responses.* Arch Pathol Lab Med, 2000. **124**(9): p. 1382-5.
11. Conner, E.R. and S.S. Saini, *The immunoglobulin E receptor: expression and regulation.* Curr Allergy Asthma Rep, 2005. **5**(3): p. 191-6.
12. Turner, H. and J.P. Kinet, *Signalling through the high-affinity IgE receptor Fc epsilonRI.* Nature, 1999. **402**(6760 Suppl): p. B24-30.
13. Bohle, B., *The impact of pollen-related food allergens on pollen allergy.* Allergy, 2007. **62**(1): p. 3-10.
14. Hiller, R., et al., *Microarrayed allergen molecules: diagnostic gatekeepers for allergy treatment.* FASEB J, 2002. **16**(3): p. 414-6.
15. Aydin, O., et al., *Omalizumab in the treatment of allergic bronchopulmonary aspergillosis: One center's experience with 14 cases.* Allergy Asthma Proc, 2015. **36**(6): p. 493-500.
16. Moverare, R., et al., *Different IgE reactivity profiles in birch pollen-sensitive patients from six European populations revealed by recombinant allergens: an imprint of local sensitization.* Int Arch Allergy Immunol, 2002. **128**(4): p. 325-35.
17. Radauer, C., P. Lackner, and H. Breiteneder, *The Bet v 1 fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands.* BMC Evol Biol, 2008. **8**: p. 286.
18. Ebner, C., et al., *Identification of allergens in fruits and vegetables: IgE cross-reactivities with the important birch pollen allergens Bet v 1 and Bet v 2 (birch profilin).* J Allergy Clin Immunol, 1995. **95**(5 Pt 1): p. 962-9.
19. Gajhede, M., et al., *X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy.* Nat Struct Biol, 1996. **3**(12): p. 1040-5.

20. Kofler, S., et al., *Crystallographically mapped ligand binding differs in high and low IgE binding isoforms of birch pollen allergen bet v 1.* J Mol Biol, 2012. **422**(1): p. 109-23.
21. Grutsch, S., et al., *Ligand binding modulates the structural dynamics and compactness of the major birch pollen allergen.* Biophys J, 2014. **107**(12): p. 2972-81.
22. Breiteneder, H., et al., *Molecular characterization of Api g 1, the major allergen of celery (Apium graveolens), and its immunological and structural relationships to a group of 17-kDa tree pollen allergens.* Eur J Biochem, 1995. **233**(2): p. 484-9.
23. Hoffmann-Sommergruber, K., et al., *IgE reactivity to Api g 1, a major celery allergen, in a Central European population is based on primary sensitization by Bet v 1.* J Allergy Clin Immunol, 1999. **104**(2 Pt 1): p. 478-84.
24. Schirmer, T., et al., *Crystal structure of the major celery allergen Api g 1: molecular analysis of cross-reactivity.* J Mol Biol, 2005. **351**(5): p. 1101-9.
25. Gepp, B., et al., *Chimeras of Bet v 1 and Api g 1 reveal heterogeneous IgE responses in patients with birch pollen allergy.* J Allergy Clin Immunol, 2014. **134**(1): p. 188-94.
26. Wilkins, M.R., et al., *Protein identification and analysis tools in the ExPASy server.* Methods Mol Biol, 1999. **112**: p. 531-52.
27. Kelley, L.A., et al., *The Phyre2 web portal for protein modeling, prediction and analysis.* Nat Protoc, 2015. **10**(6): p. 845-58.
28. Verdino, P. and W. Keller, *Circular dichroism analysis of allergens.* Methods, 2004. **32**(3): p. 241-8.
29. Whitmore, L. and B.A. Wallace, *Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases.* Biopolymers, 2008. **89**(5): p. 392-400.
30. Compton, L.A. and W.C. Johnson, Jr., *Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication.* Anal Biochem, 1986. **155**(1): p. 155-67.
31. Sreerama, N. and R.W. Woody, *Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set.* Anal Biochem, 2000. **287**(2): p. 252-60.
32. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.
33. Battye, T.G., et al., *iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 271-81.
34. Winn, M.D., et al., *Overview of the CCP4 suite and current developments.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.
35. Emsley, P., et al., *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
36. Mertens, H.D. and D.I. Svergun, *Structural characterization of proteins and complexes using small-angle X-ray solution scattering.* J Struct Biol, 2010. **172**(1): p. 128-41.
37. Petoukhov, M.V., et al., *New developments in the program package for small-angle scattering data analysis.* J Appl Crystallogr, 2012. **45**(Pt 2): p. 342-350.
38. Guhsl, E.E., et al., *Vig r 6, the cytokinin-specific binding protein from mung bean (Vigna radiata) sprouts, cross-reacts with Bet v 1-related allergens and binds IgE from birch pollen allergic patients' sera.* Mol Nutr Food Res, 2014. **58**(3): p. 625-34.

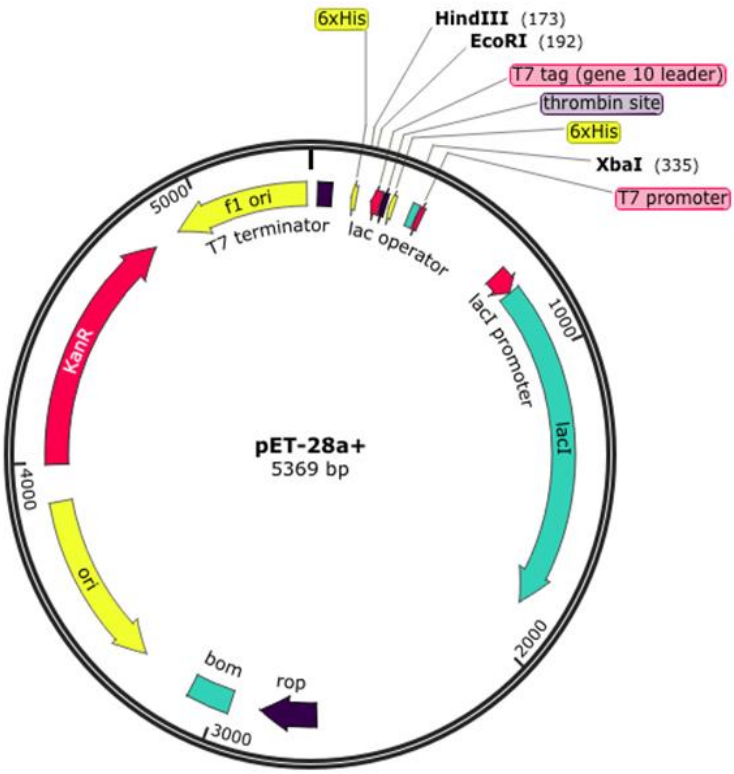
39. Wangorsch, A., et al., *Mutational epitope analysis and cross-reactivity of two isoforms of Api g 1, the major celery allergen.* Mol Immunol, 2007. **44**(10): p. 2518-27.

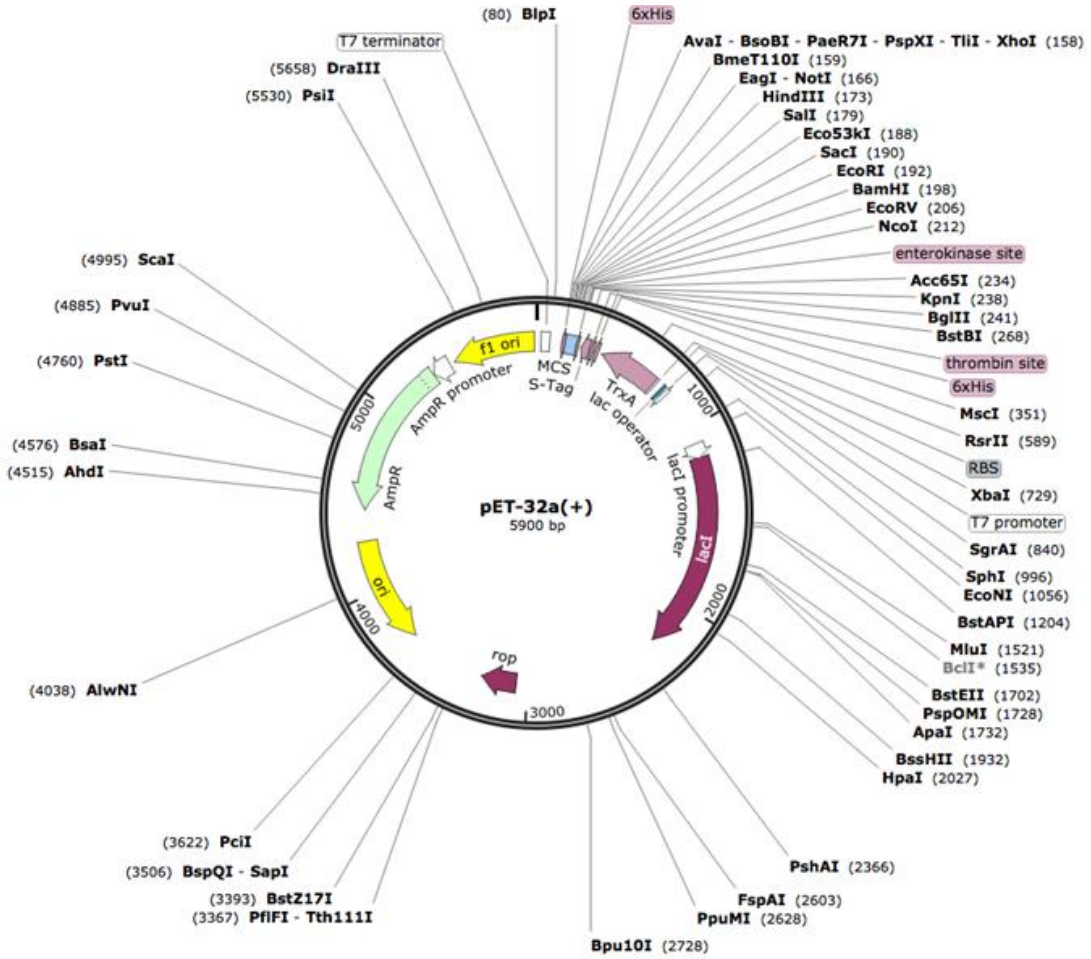
8 List of Figures

List of Figures

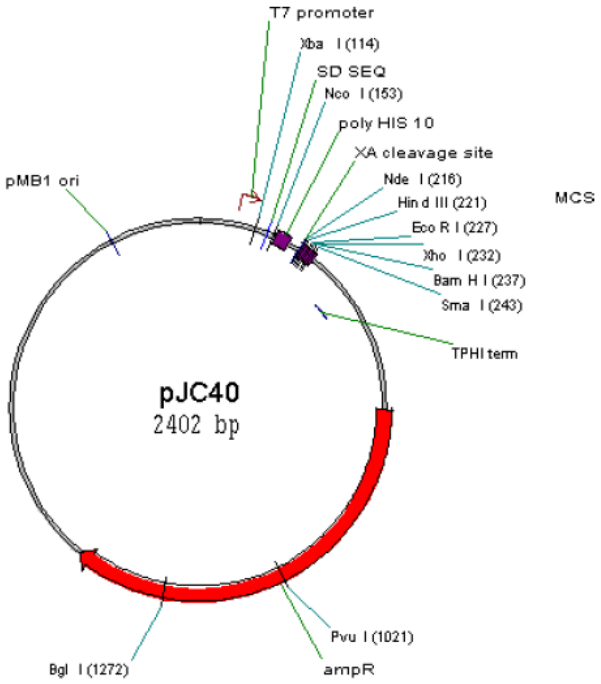
9 Supplemental

Vectors




(80) BlpI
6xHis
T7 terminator
(5658) DraIII
(5530) PsiI
AvaI - BsoBI - PaeR7I - PspXI - TliI - XhoI (158)
BmeT110I (159)
EagI - NotI (166)
HindIII (173)
SalI (179)
Eco53kI (188)
SacI (190)
EcoRI (192)
BamHI (198)
EcoRV (206)
NcoI (212)
enterokinase site
Acc65I (234)
KpnI (238)
BglII (241)
BstBI (268)
thrombin site
6xHis
MscI (351)
RsrII (589)
RBS
XbaI (729)
T7 promoter
SgrAI (840)
SphI (996)
EcoNI (1056)
BstAPI (1204)
MluI (1521)
BclI* (1535)
BstEII (1702)
PspOMI (1728)
ApaI (1732)
BssHII (1932)
HpaI (2027)
(4995) ScaI
(4885) PvuI
(4760) PstI
(4576) BsaI
(4515) AhdI
(4038) AlwNI
(3622) PciI
(3506) BspQI - SapI
(3393) BstZ17I
(3367) PflFI - Tth111I
PshAI (2366)
FspAI (2603)
PpuMI (2628)
Bpu10I (2728)
f1 ori
MCS
S-Tag
TrxA
lac operator
AmpR promoter
AmpR
ori
rop
lacI promoter
lacI
1000
2000
3000
4000
5000
pET-32a(+)
5900 bp

ATCC 87114


T7 promoter
Xba I (114)
SD SEQ
Nco I (153)
poly HIS 10
XA cleavage site
Nde I (216)
Hin d III (221)
Eco R I (227)
Xho I (232)
Bam H I (237)
Sma I (243)
TPHI term
MCS
pMB1 ori
pJC40
2402 bp
Pvu I (1021)
ampR
Bgl I (1272)

Sequencing results

Api-Bet 2

TCATTGTTTACTTTAAGAAGGAGATATACCATGGGCCATCATCATCATCATCATCATCATCACAGCAGCGG
CCATATCGAAGGTCGTCATATGGGCGTACAGACCCATGTGCTGGAACTGACTAGTGTTGTTCCGGCCGCGAAA
ATCTTCCAAGGCTTTGTCATTGACGTGGACACCGTTTTGCCGAAAGCAGCACCAGGAGCCTACAAATCGGTCG
AAATCAAAGGAGATGGTGGTCCTGGTACCCTCAAGATCATCACATTGCCCGATGGTGGTCCGATAACCACCAT
GACACTTCGCATTGATGGCGTGGATCACACGAATCTGACGTTTGACTATTCCGTGATTGACGGGGATATTCTGT
TAGGGTTCATTGAGAGCATTGAGAACCATGTCAAACTGGTGGCTACTCCAGATGGCGGCTCTATTGCGAAAAC
GACCGCAATCTTTCACACCAAAGGCGATGCCGTAGTTCCGGAAGAAAACATCAAGTACGCGAACGAACAGAAT
ACGGCTCTGTTCAAAGCGTTAGAAGCCTATCTGATTGCGAATTAAGGATCCCGGGCCCTAGCTAACTGATCCG
GCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTG
GGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATAATTCTTGAAGACGAA
AGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACT
TTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGA
CAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCT
TATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGA
GATCAGTTGGGTG

His-tag
Linker-region with Factor Xa-cleavage site
Gene Api-Bet 2
Stop-codon

ORF (ExPASy translate tool):

SLFTLRRRYTMGHHHHHHHHHHSSGHIEGRHMGVQTHVLELTSVVPAAKIFQGFVIDVDT
VLPKAAPGAYKSVEIKGDGGPGTLKIITLPDGGPITTMTLRIDGVDHTNLTFDYSVIDGD
ILLGFIESIENHVKLVATPDGGSIAKTTAIFHTKGDAVVPEENIKYANEQNTALFKALEA
YLIAN-GSRALAN-SGC-QSPKGS-VGCCHR-AITSITPWGL-TGLEGFFAERRNYIRII
LEDERAS-YAYFYRLMS---WFLRRQVALFGEMCAEPLFVYFSKYIQICIRS-DNNPDKC
FNNIEKGRV-VFNISVSPLFPFLRHFAFLFLLTQKRW-K-KMLRSVG

Api-Bet 4

TATTTACTTTAAGAAGGAGATATACCATGGGCCATCATCATCATCATCATCATCATCACAGCAGCGGCCAT
ATCGAAGGTCGTCATATGGGCGTACAGACCCATGTCCTGGAACTCACCAGTTCAGTGTCTGCCGAAAAGATCT
TTCAGGGGTTTGTGATCGATGTGGACACAGTACTGCCGAAAGCAGCTCCAGGTGCCTACAAGTCGGTTGAGAT
CAAAGGCGATGGTGGTCCTGGCACTTTGAAGATCATCACGTTACCCGATGGACTTCCGATTACCACGATGACC
CTGCGCATTGATGGGGTGAACAAAGAAGCGCTGACCTTCGACTATAGCGTGATAGACGGTGGCATTTTGGGT
GACTTCATTGAGAGCATTGAGAACCATGTCGTTCTGGTTCCGACAGCTGATGGAGGCTCCATTGCCAAAACCA
CTGCAATTTTCCACACCAAAGGCGATGCAGAAGTCAAAGCGGAACAGATCAAAGCTGCGAACGAAATGAATA
CGACGCTGTTTAAAGCGTTAGAAGCCTATCTGATTGCGAATTAAGGATCCCGGGCCCTAGCTAACTGATCCGG
CTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGG
GGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATAATTCTTGAAGACGAAA
GGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTT
TCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACA
ATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTA
TTCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAGA
TCAG

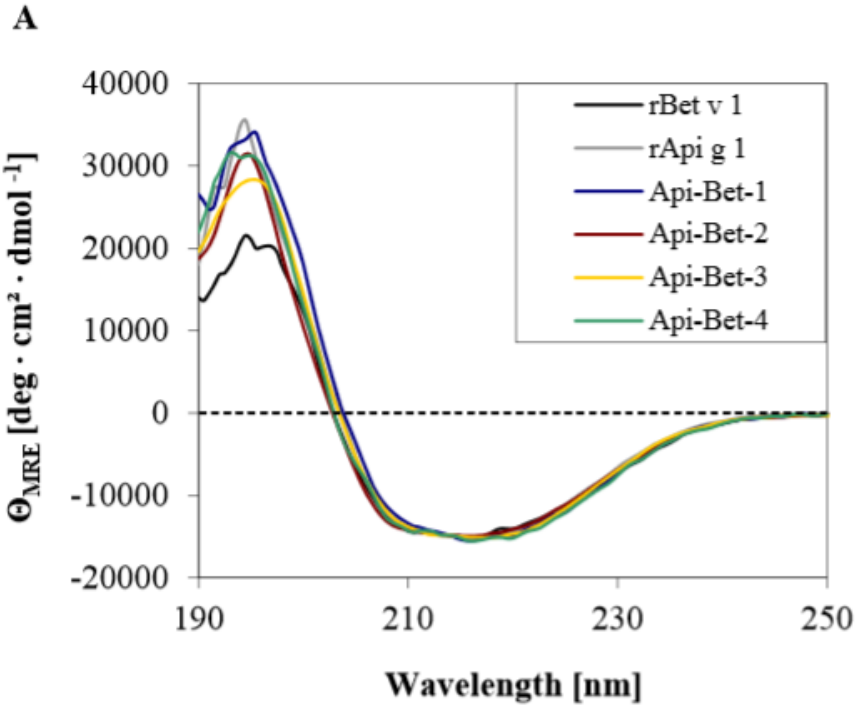
His-tag
Linker-region with Factor Xa-cleavage site
Gene Api-Bet 2
Stop-codon

ORF (ExPASy translate tool):

FTLRRRYTMGHHHHHHHHHHSSGHIEGRHMGVQTHVLELTSSVSAEKIFQGFVIDVDTVL
PKAAPGAYKSVEIKGDGGPGTLKIITLPDGLPITTMTLRIDGVNKEALTFDYSVIDGGIL
GDFIESIENHVVLVPTADGGSIAKTTAIFHTKGDAEVKAEQIKAANEMNTTLFKALEAYL
IAN-GSRALAN-SGC-QSPKGS-VGCCHR-AITSITPWGL-TGLEGFFAERRNYIRIILE
DERAS-YAYFYRLMS---WFLRRQVALFGEMCAEPLFVYFSKYIQICIRS-DNNPDKCFN
NIEKGRV-VFNISVSPLFLFCGILPSCFCSTQKRW-K-KMLRS

CD-spectra

The following figure shows the published CD-spectra of the four chimeric constructs (Api-Bet 1-4) and the parent proteins Bet v 1 and Api g 1 (recombinant) Gepp et al. [25]



The following figure shows the melting of Api-Bet 1 by overlaying of all 15 single full spectra (25°C-95°C) during the temperature scan referring to the data of chapter 4.5.