Expression of spider silk and spider silk-like proteins in potato and tobacco

Dissertation

zur Erlangung des akademischen Grades
Doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der

Naturwissenschaftlichen Fakultät I
Biowissenschaften

der Martin-Luther-Universität Halle-Wittenberg

von

Marziya Rakhimova
geboren am: 07. 09. 1970 in: Semei, Kazakhstan

Gutachter:
1. Prof. Dr. Werner Roos
2. PD Dr. Udo Conrad
3. PD Dr. Juergen Scheller

Datum der öffentlichen Verteidigung: 12. September 2007

urn:nbn:de:gbv:3-000012379
[http://nbn-resolving.de/urn/resolver.pl?urn=nbn%3Ade%3Agbv%3A3-000012379]

List of abbreviations

ADF-3	major ampullate protein from *Araneus diadematus*
ADF-4	major ampullate protein from *Araneus diadematus*
BSA	bovine serum albumin
CaMV	cauliflower mosaic virus
DNA	deoxyribonucleic acid
EDTA	ethylenediamine tetra-acetic acid
ELISA	enzyme linked immunosorbant assay
ER	endoplasmic reticulum
ECM	extra cellular matrix
ELP	elastine like polypeptide
FLAG	flagelliform silk protein
FNT	N- terminus of FLAG protein
FCT	C- terminus of FLAG protein
g	gram
h	hour
L	litre
KLH	keyhole limpet hemocyanin
LeB4SP	legumin B4 signal sequence
MaSp1	major ampullate spidroin 1
MaSp2	major ampullate spidroin 2
MS	Murashige-Skoog
m	milli
M	molar
µ	micro
min	minute
mol	mole
n	nano
OD	optical density
Ori	origin of replication
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
Poly A	polyadenylation signal
RNA	ribonucleic acid
RT	room temperature
RS	rabbit serum
sp	signal peptide
scFv	single-chain variable fragment
SDS	sodium dodecyl sulphate
SEM	scanning electron microscopy
T-DNA	transferred DNA
TSP	total soluble protein
Tw	Tween20

Contents

1 Introduction

1.1 Mechanical properties of spider silks

1.1.1 In spiders each silk is designed for a specific purpose

The term "silk" is generally associated with the protein fibre produced by the larvae of the silkworm, *Bombyx mori*. The material is drawn in continuous strands of considerable length from glands opening beneath the larva`s head and is used to construct the larval cocoon. Over 4000 years ago it was discovered in China that this protein fibre could be unwound from the cocoon and used to weave fabrics that were sheer, yet strong. This property of *Bombyx mori* silk has resulted in the formation of a large textile industry. Secondary consequences have included the complete domestication of the silkworm and considerable study into the physical and chemical properties of *Bombyx mori* silk fibres. Despite the economic importance of *Bombyx mori* silk, it is far from being the only such protein fibre produced by drawing a viscous liquid through a fine orifice as silk, regardless of its precise chemical composition or physical properties. Silks form a functionally diverse lot. They are widely used (by moths, for example) in the construction of cocoons and egg cases where they provide physical support, protection, and thermal insulation. Further examples include burrowing spiders, which use silk fibres as structural elements in the construction of their burrows and some tropical ants use the silk of their larvae to fashion leaves into nests. In many of these animals each silk designed for a specific purpose is produced by a separate gland and each has its own particular chemical structure and physical properties (Denny 1976). Perhaps the best examples of this specialisation are the spiders that spin orb-webs *Nephila clavipes,* a common orb-weaving spider, produces seven separate types of silk, which are summarised in Table 1. The silks generated by this species of spider differ from the silk produced by the domesticated silkworm, *Bombyx mori.* In the silkworm silk is primarily produced during a specific stage in the lifecycle of the organism.

Table 1. Summary of silks, their functions, and glands of origin for the golden orb weaver *Nephila clavipes.*

Type of silk	Gland	Spinneret	Function of the silk
Dragline	Major ampullate	Anterior	Orb web frame, radii, safety line
Viscid	Flagelliform	Posterior	Prey capture, core fibers of adhesive spiral
Glue-like	Aggregate	Anterior and Posterior	Prey capture, adhesive silk of spiral
Minor ampullate	Minor ampullate	Medial	Orb web frame reinforcement
Cocoon	Tubuliform	Posterior	Reproduction
Wraping	Aciniform	Anterior	Attachment disk and joining fibers
Attachment	Pyriform	Anterior	Attachment disk and joining fibers

Macromolecular structure and chemical composition determine the mechanical properties of the silks.

1.1.2 Mechanical properties of spider silks

Silk proteins, either from spiders or from insects, are especially interesting as new basic materials for the production of silk fibers with mechanical properties superior to chemical fibres and, last not least, as a renewable resource in future.
Spider silks form an interesting family of fibrous proteins in which each silk has distinct mechanical properties tailored for specific functions. The various silks have several functions, for example in reproduction and prey capture.
Spider silks are particularly remarkable materials because of their high tensile strength that is comparable to Kevlar (Table 2) in combination with a significant elasticity.

Table 2. Mechanical properties of several spider silks from *Nephila clavipes* compared with other natural and synthetic materials.

Material	**Elasticity (%)**	**Tensile strength Nm^{-2}**	**Energy to break Jkg^{-2}**
Major ampullate Dragline silk	35	$4x10^9$	$1x10^5$
Flagelliform silk	>200	$1x10^9$	$1x10^5$
Rubber	600	$11x10^6$	$8x10^4$
Kevlar	5	$4x10^9$	$3x10^4$
Steel	<1	$2x10^9$	$1x10^3$

Source: Tirell, 1996

The dragline spider silk from the golden orb weaver *Nephila clavipes* displays impressive toughness, and a balance of stiffness, strength and extensibility to reflect the native function of the silk in orb web construction and capture of flying insects (Gosline et al. 1984; Gosline et al. 1986; Cunniff et al. 1994).
The capture silk of orb-weaving spiders is used to entrap flying preys (Lorraine et al. 1995).
Capture silk fibers provide an interesting contrast in material properties compared to dragline silk fibers. The capture silk has a tensile strength that is comparable to steel, (Gosline et al. 1999), and it is extremely elastic, with the ability to be stretched to almost 10 times its relaxed contour length without breaking (Gosline et al. 1999), although the toughness is nearly the same for both types of silk fibers.
Spider silks are insoluble in most solvents, including water, dilute acid and alkali, and their resistance to digestion from most proteolytic enzymes relates to the environmental stability of these proteins in fibre form. This feature is a plus in terms of the native requirements for the webs, and also a challenge during the study of the proteins present in the web silk.

1.1.3 The general features of the gland's morphology and its product

The dragline silks of spiders are produced by the ampullate glands. The major ampullate silk consists entirely of protein, and several reports on its amino acid compositions have been published (Lucas et al. 1964; Andersen 1970; Tillinghast et al. 1984). The general features of the gland's morphology and its product were reported (Ortiz et al. 2000; Candelas et al. 1990; Candelas et al. 1986; Rodriguez et al. 1995). Candelas and co-workers showed that the major ampullate gland of *Nephila clavipes* produces a single protein. They used electrophoresis to obtain an estimated molecular weight of 320 kDa.
The histochemical, histological and ultrastructural evidence available, not only from Araneids but from a variety of representatives of the liphistiomorph, mygalomorph and araneomorph suborders, argues strongly for many silks being composed of more than one protein.
Evidence obtained by other methods from luminal contents, the silks themselves and, recently, from nucleic acid sequence data, leads to the same conclusion (Sponner et al. 2004; Xu et al. 1990; Hinman et al. 1992; Guerette et al. 1996).
Another set of glands, the flagelliforms, was described originally by Sekiguchi in 1952 and later by Peters in 1955. Andersen in the 70's conducted studies on the amino acid composition of the silks of several species of *Nephila* and *Araneus* and reported that the flagelliform's fibroin is rich in glycine and proline, both of which account for approximately 64% of the protein's amino acid residues (Andersen 1970). Rodriguez and co-workers (Rodriguez et al. 1995) showed by the electrophoretic separation of the solubilised content of the flagelliform glands an uppermost diffuse band of high molecular size, preceded by a stepladder of well-defined peptides, which have been shown to be products of discontinuous translation in three other sets of glands. Translational pausing during protein synthesis resulted in heterogeneous protein populations in both *Nephila clavipes* (Candelas et al. 1983) and *Bombyx mori* (Lizardi et al. 1979).
The steps involved in the processing of silk, have been examined at various levels (Peakall 1966; Candelas et al. 1981; Willcox 1996; Vollrath et al. 1998; Vollrath and Knight 2001).
Cytological aspects of spider silk elaboration have been studied well in the case of the ampullate glands of *Araneus* (Peakall 1965; Bell and Peakall 1969; Kovoor and Zylberberg 1972). In the major ampullate and minor ampullate glands of Araneids, and to lesser extent in the flagelliform glands, two morphological regions composing the body of the gland are recognised, the ampula and the tail. The distinction between the two is essentially based on width, the tail being a relatively narrow tube and the ampula, a sac with an expanded lumen in which the secretory products of the epithelial cells are amassed.
Techniques such as cryo SEM (scanning electron microscopy) and TEM (transmission electron microscopy) have been used to examine different regions of the *Nephila edulis* gland and duct (Vollrath and Knight 1999).
In *Nephila* golden silk spiders the secretory part of the gland is bisected into two fairly distinct zones: the A-zone in the tail of the gland and the B-zone, which extends from the paunch of the sac to the funnel. The gland eventually leads to an S-shaped spinning duct, which connects to the exit spigot for the spider silk (Vollrath et al. 2005).

The duct has been subject of detailed ultractructural studies in several types' of silk glands from *Araneus diadematus* and *Nephila clavata* (Kovoor et al. 1979; Moon et al. 1989; Moon et al. 2006).
In Araneids, the secretory epithelia of minor ampullate, flagelliform and pyriform glands are composed of two, sharply delineated regions (cell types) which exocytose histochemically and ultrastructurally distinct secretory granules into the lumen, while major ampullate glands contain either two or three regions, depending on the genus (Tillinghast 1994).
The epithelium presents a simple columnar arrangement, consisting of a single type of secretory cell, which in turn contains a single type of secretory granule. These cells are characterised by their large oval-shaped nuclei, well-developed rough endoplasmic reticulum and richness in secretory granules.
Pairs of major ampullate glands have suitably and proficiently served as a model system for the analysis of the synthesis of a secretory protein and its control (Candelas et al. 1981; Candelas et al. 1990).
Several studies have been devoted to examine changes that occur in the secretory epithelium of major ampullate gland tails following stimulation by mechanical depletion method (Plazaola et al. 1990; Plazaola et al. 1991).

1.2 Genetic control of polymer design

1.2.1 Fundamental aspects of spider silk genetics

The predatory nature of spiders and the relatively low levels of production of these silks when compared to silkworm cocoon silk have prevented domestication of spiders for the production of larger amounts of silk. Therefore, many research groups have turned to genetic engineering to construct, clone and express native or synthetic genes encoding recombinant spider silk proteins. This strategy has provided new insights into fundamental aspects of spider silk genetics, silk protein structure and function, and processing of fibrous protein materials.
The partial sequences of cDNAs and the structure of the genes encoding spider and silkworm silk fibroins from different species have been determined (Yamaguchi et al. 1989; Zhou et al. 2000; for review seeHu et al. 2006). The primary amino acid sequences deduced from partial cDNA clones showed the repetitive structure of these proteins. All these proteins have a modular nature. The repetitive portions of most characterised silk proteins from spiders can be generalised as typical sets of consensus repeats containing six types of amino acid motifs as poly-A, poly-(GA), GGX, GPGXX, GPX and spacers, were (X is occupied by a limited subset of amino acids: Ala, Ser, Tyr, and Val). The repetitive amino acid motifs of the different silk proteins are thought to be directly responsible for the mechanical properties of spider silk fibers (Hayashi et al. 1999).
The major ampullate spider silk protein MaSp1 contains glycine rich three amino acid motifs and glycine-alanine stretches whereas the MaSp2 protein contains pentapeptide motifs.

A motif of amino acid sequence, GPGGX, is repeated frequently in MaSp2 dragline silk protein and flagelliform silk (core fiber of capture silk) from the golden orb-weaver *Nephila clavipes* (Hayashi et al. 1998).
MaSp1 and 2 proteins consist of repetitive sequences containing polyalanine regions, sandwiched between sequences rich in glycine.
In contrast to dragline silk proteins flagelliform silk proteins do not contain alanine stretches, however they have proline rich pentapeptide motifs, which occur in tandemly arrayed repeats up to 63 times followed by the three amino acid motif GGX. FLAG comprises also a spacer motif, which is 28 amino acid long (Table 3).

Table 3. Consensus amino acid repeats for dragline silk and flagelliform silk proteins

Silk	**Known proteins**	**Consensus amino acid repeat**
Major ampullate dragline	MaSp1*	GGAGQGGYGGLGGQGAGR GGLGGQ$(GA)_2$ A_5
	MaSp2**	$(GPGGY)_{4-7}$ $GPSGPGSA_8$
Flagelliform	FLAG***	$(GPGGX)_{43-63}$ -$(GGX)_{6-12}$ -FLAG spacer

*Xu and Lewis 1990, ** Hinman and Lewis 1992, *** Hayashi and Lewis 1998

1.2.2 Molecular architecture of FLAG

Hayashi and co-workers extensively studied flagelliform silk gene by sequencing of genomic DNA and reported that spider silk protein sequences are highly hierarchically organised. First, each amino acid motif is tandemly repeated itself, each motif itself is known to form its own secondary structure. Than all three motifs are organised into ensemble repeat. In case of flagelliform silk proteins each ensemble repeat is encoded by one exon.
The *FLAG* gene has about 30 kb long transcripts, which consist of more than 13 exons, where the first two exons encode the non-repetitive amino-terminal region of FLAG, while the conserved C-terminus is encoded by the last exon. All other exons encode repetitive parts of *FLAG*. The repeated exons are of similar length (about 1320 bps) and identical organisation. The main differences among the repeated exons are the variable numbers of tandem $(GPGGX)_n$ motifs which varies within a range at 43 to 63 times and GGX motifs variations from 6 to 12 times. The iterated sequences within exons encode protein monomers. The introns separating the repeated exons also share high similarity. The introns 3 to 12 are each about 1420 bp long (Hayashi et al. 2000).
In contrast to *FLAG* other known spider silk genes are encoded by single exons of bigger size (Xu et al. 1990; Hinman et al. 1992; Beckwitt et al. 1998).

1.2.3 Non-repetitive C- and N- terminal parts of spider silk proteins

Essential features of all spider silk genes are the extended gene size, a high GC content of up to 60% and their high repetitiveness. Because of their big size all spider silk genes are partial sequences, mostly cDNA sequences starting from the C-terminus. All known spider silk gene family proteins including MaSp1, MaSp2, and FLAG end with distinct non-repetitive C-terminal regions. The C-terminal region is conserved in sequence and length even among phylogenetically distant species and among silk types (Hinman et al. 1992; Beckwitt et al. 1994; Guerette et al. 1996; Gatesy et al. 2001). In order to elucidate potential roles of the C-termini Sponner and co-workers investigated their presence in the spinning dope and the filament. Their obtained results showed that C-termini of dragline silk proteins MaSp1 and MaSp2 are *bona fide* parts of mature spidroins (Sponner et al. 2005). It was speculated that the C-termini of spidroins might play a crucial role in the formation of the insoluble filament in analogy to the situation found in the silk worm *Bombyx mori* (Jin and Kaplan 2003). Huemmerich and co-workers reported that repetitive parts of dragline silks are more involved in determination of the solubility of spider silk proteins, whereas non-repetitive C-terminal parts have been shown to have no significant influence on this process (Huemmerich et al. 2004). Also, their high conservation may reflect an important function that might be regulatory rather than influencing the structure. Additionally the carboxy-terminal peptides produced in the major ampullate gland and flagelliform silk produced in the flagelliform gland contain one or two cysteine residues that might serve to cross-link the spider silk fibroins by the formation of disulphide bridges (Craig et al. 1997; Sponner et al. 2004; Mello et al. 1994).
Due to a high GC content, repetitiveness and big sizes of spider silk genes, most published spider silk sequences lack the 5'- end of the mRNA. This is probably due to technical cloning circumstances that prefer amplification of 3'-regions. Despite these difficulties, extensive sequences of Flagelliform silk genomic DNAs and cDNAs of *Nephila clavipes* and *Nephila madagascariensis* for the first time provided N-terminal non-repetitive sequences of these proteins (Hayashi et al. 1998; Hayashi et al. 2000).
During the past decade also N-terminal domains have been characterised for cylindriform and for major ampullate silk protein (Motriuk-Smith et al. 2005; Rising et al. 2006). In addition to these N-terminal domains, two N-terminal sequences, encoding egg case protein 1 and 2 (ECP-1 and -2) have been reported (Hu et al. 2005; Hu et al. 2006).

1.3 Expression of silk proteins in heterologous systems

1.3.1 Silk protein polymer synthesis in prokaryotic and eukaryotic hosts

More and more scientists have attempted to create spider's webs independently of spider itself through genetic engineering by manufacturing the proteins, which constitute the silk fibers of the webs, through the use of bacteria, yeast, plants, and mammalian cells in tissue culture.
Many research groups have utilised this repetitive gene structure to genetically engineer spider silk proteins using synthetic oligonucleotide versions of the consensus repeats of

variants of these repeats (Prince et al. 1995; Lewis et al. 1996; Fahnestock et al. 1997; Scheller et al. 2001).
From partial cDNA sequences from *Nephila clavipes* encoding the proteins MaSp1 and MaSp2 a consensus repeat from MaSp1 was identified [GGAGQGGYGGLGSQGAGRGGLGGQGAG], followed by a polyalanine region (responsible for β-sheet formation) (Xu et al. 1990)
For MaSp2, the consensus repeat sequence [GPGGYGPGQQGPGGYAPGQQPSGPGS] was also followed by a short polyalanine region (Hinmann et al. 1992). However, there is less sequence conservation between the repeats of MaSp2 when compared to MaSp1 (Hayashi et al. 1999; Foo et al. 2002).
The consensus sequences for MaSp1 and MaSp2 were used to prepare synthetic genes that were subsequently multimerised through head-to-tail ligation leading to different lengths of the artificial genes (Prince et al. 1995). There are reports including variations on the synthetic gene construction strategy as well as the use of either *Escherichia coli* or the yeast *Pichia pastoris* as host systems for the expression of the heterologous proteins (Fahnestock and Bedzyk 1997). The most extensively used host system for the expression of such genes is *Escherichia coli*.
Fahnestock and Irwin designed synthetic genes to encode analogs of MaSp1 and MaSp2 from *Nephila clavipes* dragline silk. Various sizes and types of sequences were cloned and expressed. Yield and homogeneity of the products of longer genes were limited by premature termination during synthesis. Protein purification was based on the nickel affinity separation as well as ion-exchange chromatography (Fahnestock and Irwin 1997).
Fahnestock and Bedzyk used the methylotrophic yeast *Pichia pastoris* to synthesise spider dragline silk-like polymers. *Pichia pastoris* is a useful host for the production of heterologous proteins since it has been developed for large-scale fermentation (Cregg et al. 1985). The synthetic genes were expressed at high levels under the control of a methanol-inducible promoter and a histidine tag was used for the purification based on to Ni/NTA/agarose columns.
The yeast system was compatible with gene stability since large clones were successfully expressed with no apparent deletions.
Aside from the extensive studies on synthetic gene designs to study dragline silk, silk proteins from other families are studied now in a similar fashion. Zhou and co-workers described the biosynthesis of a recombinant polypeptide that mimics the dominant repeat sequence [Gly-Pro-Gly-Gly-Xaa] of the *Nephila clavipes* flagelliform silk protein in *Escherichia coli*. The authors made structural investigations of this model polypeptide providing insight into the conformation of the polypeptide in solution and the solid state. The spectroscopic information suggests that the pentapeptide segments of the flagelliform silk protein adopt a β-turn conformation in the fiber and that the mechanism of elasticity may resemble that proposed for other β-turn forming polypeptides including elastin (Zhou et al. 2001).
In general, expression levels obtained from synthetic silk genes are low, with yields of about 1 to 10 mg/L, representing usually less than 5% of the total protein in the cell, depending upon the size of the protein. The exception to these low yields appears to be the yeast system. It should also be noted that while expression levels of 300 to 1000 mg/L of protein have been reported from this host, the level of purity is not always clear.

Precipitation and non-specific interactions are also reasons for significant losses of protein during the purification process.
Both native and genetically engineered versions of spider dragline silk produced in microorganisms tend to self-assemble into microfibrills, causing precipitation and leading to loss of protein during purification (Kaplan et al. 1992; Wilson et al. 2000). Recently in the group of Dr. Th. Scheibel, TU Muennich, bacterial expression system for spider silk genes have been developed (Scheibel 2004). Recent efforts have focused on alternative methods to maintain solubility of the recombinant versions of these proteins by alterations of the native or consensus sequences to include encoded "triggers" to regulate the molecular-level assembly of these proteins (Winkler et al. 1999; Szela et al. 2000; Wilson et al. 2000; Winkler et al. 2000). The low expression levels of spider silk proteins in fast-growing microorganisms such as yeast and bacteria have prompted studies into alternative hosts.
A recent report by Lazaris and co-workers demonstrated that partial cDNA clones from *Araneus diadematus* dragline protein could be heterologously expressed and isolated from mammalian epithelial cells and subsequently spun into fibers with mechanical properties close to those of native silk fibers (Lazaris et al. 2002). Since this report, follow-up papers are lacking as well as reports about further development of this technology.

Molecular farming in transgenic plants opened the way to set up production systems with the theoretical capacity of large scale production.

1.3.2 Molecular farming of proteins

"Molecular farming" describes the application of molecular biological techniques to produce proteins, carbohydrates and secondary metabolites in transgenic plants for commercial purposes. Many of the proteins being farmed in plants are biopharmaceuticals as therapeutic antibodies or vaccines and aimed at improving human and animal health (Fischer et al. 2004; Stoger et al. 2005). In the area of molecular farming many exciting and potentially beneficial developments in plants biotechnology are taking place. Generally, also materials for technical use can be produced in plants.
However, the production of a complex protein that is folded, processed, easily purified and safe for pharmaceutical use, all at low cost, is a real challenge.

1.3.2.1 Advantages of plant production system

Plant production of recombinant proteins offer several advantages. Protein synthesis, secretion and post-translational modifications are similar in plant and animal cells, there being only minor differences in protein glycosylation (Fischer et al. 2000). Slight differences in codon usage in plants can be compensated by adjusting transgene sequences allowing them to correctly assemble mammalian multimeric proteins (Fischer et al. 2000). In addition, products from transgenic plants are unlikely to be contaminated by animal pathogens, microbial toxins or oncogenic sequences (Fischer et al. 2000).

Mammalian cell systems are expensive and cannot easily be scaled up. Bacterial systems can be scaled up, but often the recombinant proteins are not properly processed (they are not properly folded and disulphide bridges are not formed), so intracellular precipitation of non-functional proteins can occur. Plant systems can be scaled up so that large amounts of material can be harvested and processed, allowing industrial scale amounts of protein to be purified.

1.3.2.2 Transgenic plants are powerful protein factories

In general all fibrous proteins of spiders and insects are large (230-720 kDa) and have highly repetitive amino-acid sequences.
Nevertheless, the genetically programmed synthesis of protein polymers allows precise control of molecular mass and amino acid monomer composition, and therefore could lead to the design and synthesis of polymers whose differing physical and functional properties are exactly controlled. Such controlled synthesis is difficult using chemical technologies (O'Brien 1993).
It seems to be, that transgenic plants are powerful protein factories that can overcome the problems mentioned above. There are several examples of stably transformed plants expressing various spider-silk genes summarised in Table 4.

Table 4. Spider silk proteins produced in plants.

Protein	**Promoter**	**Species**	**Remarks**	**Size of protein**	**Reference**
Synthetic spider silk proteins like MaSp1	35S CaMV	Tobacco, ER	2% (TSP)	14-120 kDa	Scheller et al. 2001
Synthetic spider silk proteins like MaSp1	35S CaMV β-conglycinin α' subunit	Arabidopsis, Leaves Seeds Soybean, Somatic embryos	1.65%(TSP) 0.06%(TSP) 1.4% (TSP) 1% (TSP)	64 kDa 127 kDa 64 kDa 127 kDa 64 kDa	Barr et al. 2004
Native MaSp1 and Masp2	35S CaMV	Tobacco (field trials)	0.1% (TSP) 69 g per hectare	60.3 kDa and 58.6kDa	Menassa et al. 2004
Synthetic spider silk proteins like MaSp1(SO1-100xELP)	35S CaMV	Tobacco, Leaves Potato	Up to 4%(TSP) 80mg per kg	94.4kDa 94.4kDa	Scheller et al. 2004

Scheller and co-workers constructed stable transgenic tobacco and potato lines to express various synthetic spider-silk genes ranging from 420 to 3600 base pairs in size. The genes were assembled so as to achieve very high homology to the native MaSp1 gene from *Nephila clavipes* (more than 90%). Accumulation of up to 2% spider-silk protein of total soluble protein (TSP) in the endoplasmic reticulum (ER) of tobacco leaves was achieved, and this production was relatively independent of the size of the spider-silk protein. The recombinant spidroins exhibited extreme heat stability. This property, together with resistance against acidification and fractionated ammonium sulphate precipitation, was used to purify the plant produced spidroins by a simple and efficient procedure (Scheller et al. 2001).
Barr and co-workers showed the expression of synthetic spider silk proteins in *Arabidopsis thaliana* driven by the ubiquitous CaMV 35S promoter and by the seed-specific β-conglycinin α' subunit promoter. In *Arabidopsis* leaves, a 64 kDa artificial silk protein mimicking MaSp1 was accumulated to 1.65% TSP and a 127 kDa protein to 0.06% TSP (Barr et al. 2004).
In further experiments, Scheller and co-workers constructed a fusion protein of 94.4 kDa made of the synthetic spider silk protein SO1 (51.2 kDa) and the synthetic elastic biopolymer 100xELP (100 repeats of pentapeptide Val-Pro_Gly_Xaa_Gly [where Xaa is Gly,Val or Ala] (Meyer and Chilkoti 1999).
The 'best- producer' plants accumulated up to 4% of TSP as spider-silk-elastin fusion proteins (Scheller et al. 2004).
Natural elastin fibers provide elasticity to many tissues that require the ability to be deformed repetitively and reversibly (Rosenbloom et al. 1993). Even when expressed as a fusion protein, elastin-like polypeptides become reversibly insoluble if the temperature is raised above their transition temperature (Meyer and Chilkoti 1999), when the temperature decreases, elastine-like polypeptides and their fusion proteins become soluble again ('inverse transition cycling') (Urry et al. 1988).

1.3.2.3 Purification and possible medical application of recombinant spider silk-like proteins

A first step towards the production of new biomaterials that have useful industrial and medical properties has been taken by combining a spider silk protein that exhibits a high tensile strength with an elastic biopolymer, such as an elastin-like polypeptide.
A new purification strategy from plants has been adapted using the unique properties of ELP fusions and the techniques developed for bacterial systems (Scheller et al. 2004; Meyer and Chilkoti 1999). First, the spider silk proteins from plants were enriched by boiling. The authors used heat treatment at 95°C for 60 min and clearance by centrifugation of tobacco leaf extracts to enrich the spider-silk-ELP fusion protein in the supernatant. For the selective precipitation of the SO1-100xELP, NaCl was added to a final concentration of 2 M and the temperature was raised to 60°C. Under these conditions, the recombinant spider-silk-elastin fusion proteins aggregated and could be precipitated by centrifugation. Cellular proteins remained in the supernatant. The precipitated recombinant proteins were resolved at a lower temperature and without salt to a final concentration of 1 mg/mL of nearly homogenous product. Dialysis against

water and drying led to the formation of silk-proteins in the form of storable membranes. Extraction of 1 kg tobacco leaf resulted in 80 mg pure recombinant spider-silk-elastin fusion proteins (Scheller et al. 2004).
The selection of suitable spidroin and fibroin sequences for defined applications and commercialisation is still at an early stage. To date, there are no reports of the successful spinning of plant-derived silk proteins. The first example of possible medical use in future has, however, been reported (Scheller et al. 2004).
Adherent mammalian cells need extra cellular matrix (ECM) proteins for attachment, proliferation and differentiation *in vivo* as well as *in vitro*. For successful medical application, recombinant ECM-like proteins should enhance cell growth in cell culture, inhibit differentiation and exhibit a high biocompatibility.
The production of such proteins in plants will avoid contamination with mammalian viruses. In a first attempt to test spider-silk-ELP fusion proteins as substitutes for original ECM proteins, the growth of human chondrocytes (HCH-371) and CHO cell lines on this plant derived fusion protein were compared with the growth of these cells on fetal calf serum and collagen. The growth of the cells on spider-silk-ELP was comparable to that on collagen.
The plant-derived spider-silk-elastin fusion protein seems to be an effective, biocompatible matrix that promotes the growth of mammalian cells of various characters. In addition, mechanical properties of membranes derived from plant-produced spider silk-ELP have firstly been estimated (Junghans et al. 2006).

1.4 Research objectives

Since first partial cDNA sequences are obtained for spider dragline silk proteins, the numbers of recombinant studies are remarkably increased in this field. But, nevertheless, structures of spider silk proteins, which belong to the fibrillar proteins group, still need to be investigated. In contrast, the structure of globular proteins is well studied. The investigations of structure-function relationship in fibrous proteins are a challenge of modern molecular biology.
Spider silk proteins can serve as the best models for structure-function relationship studies due to their iterated sequences and wide range of spider silk proteins available even from single spider. Each silk consist of its own unique combination of structural modules.
Recently biophysical studies on native spider silk proteins revealed that elasticity of spider silk proteins is due to secondary structure formed by the pentapeptide motif, while alanine or glycine-alanine stretches form crystalline β-sheet structures which is tailored to tensile strength (Hayashi et al. 1999; Zhou et al. 2001).
Among all spider silk proteins FLAG is known to be the stretchiest silk. Studies on FLAG protein's structure-property relationships will shed light on the basis of elasticity in spider silk proteins. FLAG is one of few spider silk proteins, where N-terminal and C-terminal sequences are available for studies. Northern blot analysis showed FLAG's mRNA to be 15500 bp in length (Hayashi et al. 1998). Such an extended and highly repetitive gene is impossible to express in plants directly, therefore we aimed to catch key features of the repetitive part of deduced flagelliform sequences and additionally to add

to this key repetitive motifs N-terminal and C-terminal non-repetitive sequences. Functions of N- and C- terimines are not studied. So far, the most studied recombinant spider silk protein in different heterologous systems including plants were dragline silk proteins. Previously, recombinant studies used native and synthetic dragline genes either without N- and C-terminus (Fahnestock et al. 1997; Scheller et al. 2001) or only with C-terminus (Arcidiacono et al. 1998; Huemmerich et al. 2004). This is mainly due to the absence of N-terminal sequences of available dragline silks genes.
Previous studies on spider silk proteins expressed in plants (Scheller et al. 2001) showed that plant systems can maintain highly repetitive genes of spider silk proteins without recombination and deletion events. Also, it was demonstrated, that spider silk proteins are not toxic for plant cells, when the protein is targeted to the ER.
Spider silk proteins derived from plants were shown to be heat stable, tolerant to acidification and could be purified by ammonium sulphate precipitation.
Taking into consideration, all of the mentioned advantages of plant expression system, we decided to express FLAG protein sequences in tobacco plants. Furthermore, spider silk proteins production in tobacco can be easily scaled up, thus providing enough material for fiber spinning experiments.
FLAG's N- and C- terminies possibly provide stability to spider silk proteins in cells. Other predicted functions for the non-repetitive parts of spider silk proteins are their possible involvement in self-assembling and polymerisation of protein molecules. To verify if the C-terminus of FLAG, containing two cysteine residues is involved in polymerisation of protein monomers, *FLAG* gene variants lacking C-terminus were designed.
For the goal of our research it was important to operate with FLAG proteins without any additional peptide sequences as immunodetection tags. Therefore it was necessary to raise recombinant and conventional antibodies against the non-repetitive parts of FLAG.
Our plant expression cassette provided at the C-terminus the ER retention signal KDEL to accumulate spider silk proteins in the ER. This fusion with KDEL of target protein in many cases resulted in higer expression.
For plants (tobacco and potato) the successfully expression of spider silk proteins was demonstrated (Scheller et al. 2001). Therefore potato tubers are considered as biofactories. The agronomical infrastructure for harvesting and processing of potatoes is already available. Starch potatoes can be considered as a by-production system, where simultaneously technical starch and recombinant proteins can be produced. Therefore, one task of this work was the expression of spider silk proteins in starch potatoes.
Spider silk-ELP fusion proteins can be purified from plant crude extract based on heat stability of spider silk proteins and "inverse transition cycling" (see above).
The size of structural proteins expressed in plants is known to be in range (10-100 kDa) (Scheller and Conrad 2005) and size of spider silk in nature is shown to be in range (220-720 kDa). This big size of spider silk proteins can be achieved for production in plants *via* post-translational modification as for example by dimerisation of protein molecules *via* disulphide bridges of cysteine residues. Closure of disulphide bridges can occur in the ER. As a first trial for post-translational modification of recombinant expressed proteins tobacco and starch potato cv. Albatros were transformed with a Fc-100x ELP fusions encoding gene, where Fc is a part of human immunoglobulin G that contains cysteine residues.

So, our research on spider silk proteins in plants can be considered as input in the general knowledge about structure and properties of newly expressed plant-derived FLAG protein and as input into the development of molecular farming strategies for technical proteins.
The expression of spider silk proteins, for first time with its N-terminus and C-terminus in plants can shed light on the functional roles of the non-repetitive parts of FLAG in silk protein self-assembly.
Antibodies against N-terminus and C-terminus of FLAG could be used for the immunodetection of FLAG proteins in plant crude extracts. Also, it is possible to use them in immunolocalisation studies on spider silk proteins in plant cells or in spiders.
Expression of spider silk-ELP fusions in starch potato and preliminary analyses of transgenic potato plants under greenhouse conditions allows performing an evaluation of potato tubers as biofactory for production of spider silk-ELP fusion proteins in big scales in field trials.

2 Results

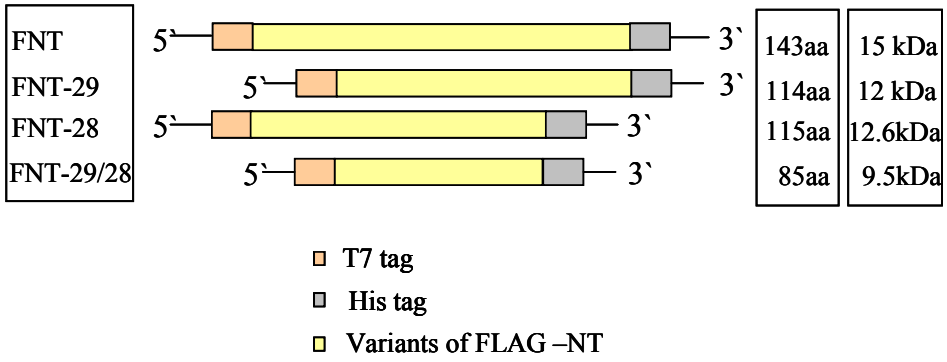
2.1 Raising antibodies against FLAG's N-terminus and C-terminus

The production of spider silk proteins in plants needs specific immunological tools to monitor the expression and to detect the spider silk protein of interest during extraction, purification and storage. The use of tags should be avoided in some cases to exclude any influence of these peptides on the spider silk protein properties. Spider silk proteins are mainly highly repetitive and therefore their immunogenecity is rather low. For some spider silk proteins non-repetitive parts are known. In the case of the major ampullate spidroins 1 and 2 (MaSp 1 and 2) from *Argiope trifasciata* and *Araneus gemmoides*, respectively the non-repetitive C terminal peptide has been cloned (Huemmerich et al. 2004). For the flagelliform spider silk protein FLAG from *Nephila clavipes*, non-repetitive N- and C-terminal peptides have been described (Hayashi et al. 1998). These peptides could serve as suitable antigens.

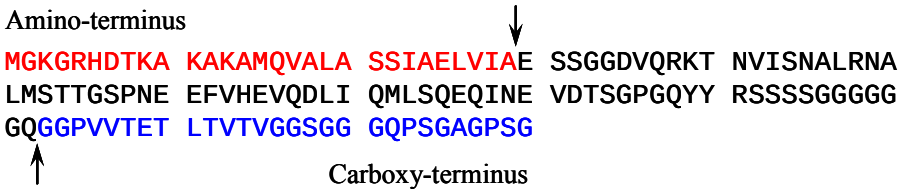
2.1.1 Cloning of non-repetitive sequences of *FLAG* and *MaSp 1* and *MaSp 2* into the expression vector pET 23a

The main object of this part of work is the isolation and characterisation of scFv antibody fragments and the production and characterisation of polyclonal rabbit sera specific either for the N-terminus or for the C-terminus of FLAG.
Therefore, N-terminal and C-terminal non-repetitive sequences of FLAG have to be expressed and purified to provide enough protein for Phage Display and for immunisation of rabbits.
The DNA sequences encoding N-terminus and C-terminus of *FLAG* are synthesised by PCR performed with FOR Flag_C-Ter, BACK Flag_C-Ter and FOR Flag_N-Ter, BACK Flag_N-ter primers and pCYH1, pCYH2 (kindly provided by Dr. C.Y. Hayashi, GenBank accession numbers AF027972 [pCYH1], AF027973 [pCYH2]) used as templates, respectively. The obtained PCR products were 273 bp and 390 bp in size for the C-terminus and the N-terminus of *FLAG*, respectively and had newly inserted XhoI and EcoRI restriction sites at their 5'- and 3'- ends. These PCR products were ligated into the pET 23a (+) vector, which was cut with XhoI and EcoRI restriction enzymes. The same cloning strategy was used for truncated N-termini versions of the FLAG protein and the C-termini of MaSp1 and MaSp2 (Figure 1A, C). The sizes of the PCR products for MaSp 1 and MaSp 2 were 270 bp for each, whereas sizes of PCR product for three truncated versions of FLAG NT were 303 bp, 310 bp and 219 bp for FNT-29, FNT-28 and FNT-29/28, respectively.

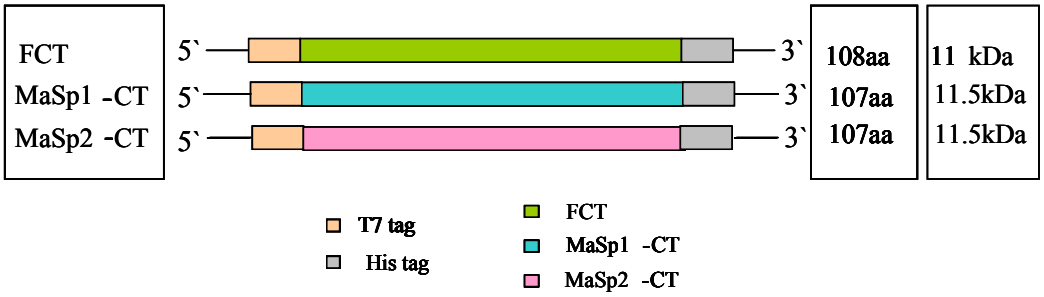
A


FNT
FNT-29
FNT-28
FNT-29/28
5`
3`
143aa
114aa
115aa
85aa
15 kDa
12 kDa
12.6kDa
9.5kDa
T7 tag
His tag
Variants of FLAG –NT

B


Amino-terminus
MGKGRHDTKA KAKAMQVALA SSIAELVIAE SSGGDVQRKT NVISNALRNA
LMSTTGSPNE EFVHEVQDLI QMLSQEQINE VDTSGPGQYY RSSSSGGGGG
GQGGPVVTET LTVTVGGSGG GQPSGAGPSG
Carboxy-terminus

C



D

```
ADF3 *(1) RGSMGAASAAVSVGGYGPQSSSAPVASAAA--------------------
ADF4 *(1) RGSMGAYGPSPSASASVAA-----------------------------

FLAG CT(1)SRVPDMVNGIMSAMQGSG--FNYQMFGNMLSQYSSGSGTCNP---N----

ADF3 (31) SRLSSPAASSRVSSAVSSLVSSG-PTNQAALSNTISSVVSQVSASNPGLS
MaSp1 (1) SRLSSPGAASRVSSAVTSLVSSGGPTNSAALSNTISNVVSQISSSNPGLS
MaSp2 (1) SRLSSPAASSRVSSTVSTLASSG-PSDASVVSSALSNLVSQVSASQPGLS
ADF4 (20) SRLSSPAASSRVSSAVSSLVSSG-PTNGAAVSGALNSLVSQISASNPGLS

FLAGCT(42)NVNVLMDALLAALHCLSNHGSSSFAPSPTPAAMSAYSNSVGRMFAY--- (87)

ADF3 (80)GCDVLVQALLEVVSALVSILGSSSIGQINYGA---------------- (112)
MaSp1(51)GCDVLVQALLEIVSALVHILGSANIGQVNSSGVGRSASIVGQSINQAFS (100)
MaSp2(50)GCDVIVQALLELVSALVHILGSSSLGQVDYNGASYSAQRLSQALGNTLG (99)
ADF4 (70)GCDALVQALLELVSALVAILSSASIGQVNVSSVSQSTQMISQALSG--- (115)
```

Figure 1. Antigens which were expressed in bacterial expression system and purified *via* affinity chromatography. A: schematic representation and sizes of FLAG NT variants; B: amino acid sequences of FLAG N-terminus, predicted 29 aa signal peptide sequence is marked in red, 28 amino acid from carboxy terminus is marked in blue. C: schematic representation of C-termini of FLAG, MaSp1 and MaSp2 and their sizes, D: alignment of amino acid sequences of C - termini of FLAG, MaSp1 and MaSp2 with C-termini of ADF-3 and ADF-4 (*Araneus diadematus* fibroin) and their sizes in aa. * ADF3 and ADF4 were kindly provided by Dr. Th. Scheibel (TU Muenchen).

2.1.2 Expression and purification of recombinant non-repetitive peptides of FLAG and MaSp 1 and MaSp 2 from bacterial cell culture under native and denaturing conditions

The used pET system is a powerful protein expression tool, because protein expression *via* the very strong T7 polymerase/T7 promoter-system can result in a very high yield of recombinant protein, and the presence of different tags, like the His tag allows the easy and efficient purification of the tagged protein.
Before deciding on a purification strategy, it is important to determine whether the protein is soluble in the cytoplasm or located in cytoplasmic inclusion bodies. Many proteins form inclusion bodies when they are expressed at high levels in bacteria, while others are well tolerated by the cell and remain in the cytoplasm in their native configuration. Soluble proteins are found in the supernatant and insoluble proteins are found in the pellet fraction.
The results of recombinant protein expression in bacterial cell culture show the accumulation of sufficient amount of recombinant protein either in the pellet (Figure 2A) or in the supernatant (Figure 2B). The full-length N-terminus of FLAG (FNT) is present in equal amounts in both fractions, whereas C-termini of spider silk proteins are present only in insoluble inclusion bodies. Interestingly, removal of the predicted signal peptide, which is supposed to be 29 amino acids in length, from the 5'- end of N-terminus of FLAG increases the solubility of this truncated version of the protein. The N-terminus of FLAG, truncated by 28 amino acids at its 3'-end (FNT-28), shows a similar expression level as the truncated FNT-29 protein and was partially soluble as shown in a Western blot analysis (Figure 3). When expressing the N-terminus of FLAG, truncated at both the 5'- and the 3'- ends (FNT-29/28) the protein was present in the supernatant only, i.e. it is fully soluble. This protein was also expressed at a higher amount than the other two truncated versions.

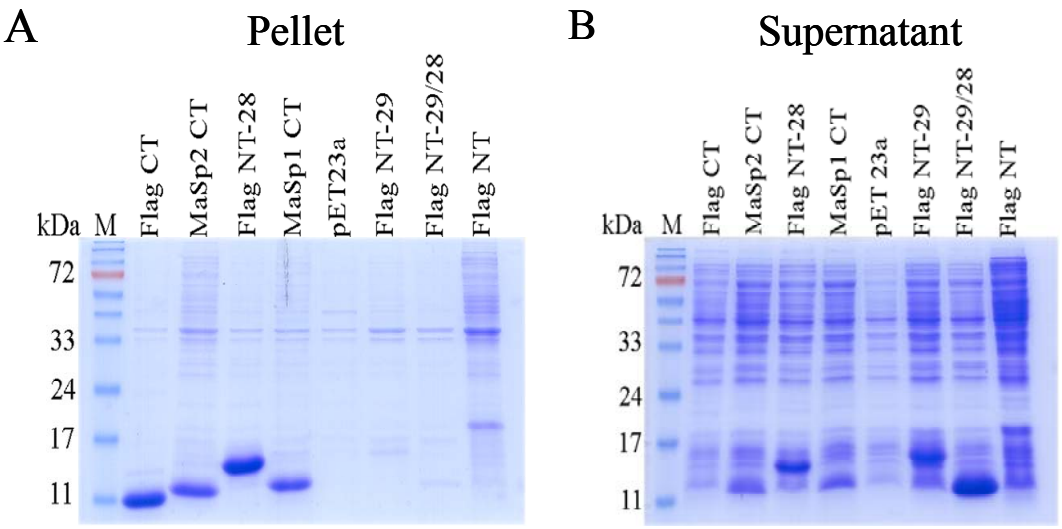


Figure 2. Analysis of N-terminal and C-terminal peptides of spider silk proteins after bacterial expression. Extracts were separated by PAGE and stained using Coomassie Blue R250. (A) insoluble proteins, (B) soluble proteins of FCT –C-terminus of FLAG, M1CT-C-terminus of MaSp1, M2CT-C-terminus of MaSp2, pET23a-control, FNT-N-terminus of FLAG, FNT-28, FNT-29, FNT-29/28- truncated versions of FLAG's N-terminus.

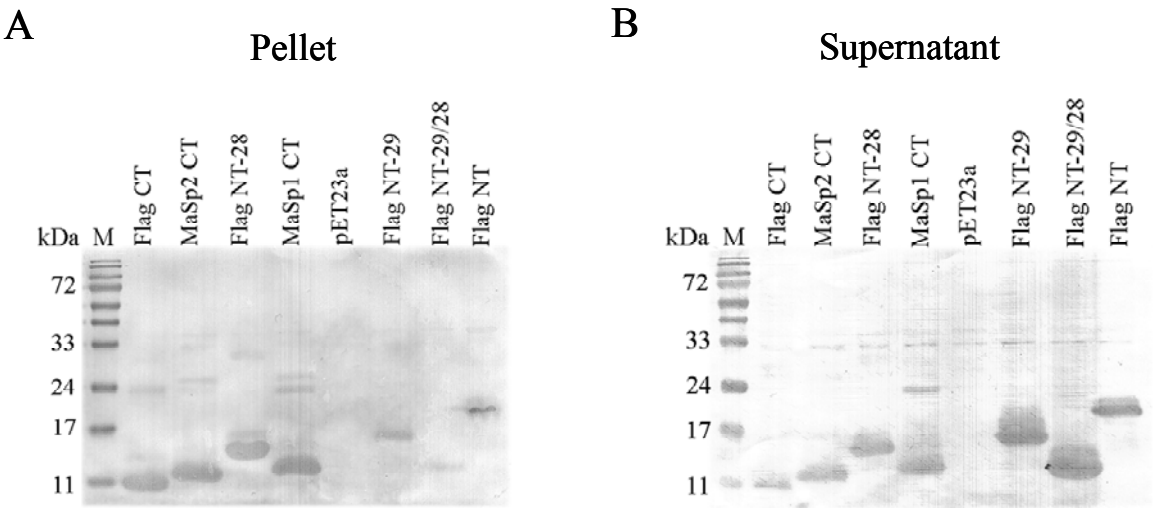


Figure 3. Western blot analysis of N-terminal and C-terminal peptides of spider silk proteins after bacterial expression. Extracts were separated by PAGE, blotted and analysed using the anti T7 tag system. (A) insoluble proteins, (B) soluble proteins of FCT: C-terminus of FLAG; M1CT: C-terminus of MaSp1; M2CT: C-terminus of MaSp2; pET23a-control; FNT: N-terminus of FLAG; FNT28, FNT29, FNT29/28: truncated versions of FLAG's N-terminus.

To facilitate rapid and efficient purification, an affinity chromatography purification system was used, which employs a metal chelate adsorbent- Ni-NTA resin. The power of

this system is based on the remarkable affinity of Ni-NTA resin for proteins and peptides that contain six consecutive histidine residues- the 6xHis affinity tag- at either their N- or C-terminus (Hochuli et al. 1988).
Most insoluble proteins form inclusion bodies in *Escherichia coli* which can be easily solubilised in denaturants such as 6 M guanidine hydrochloride or 8 M urea. These denaturants, as well as many detergents, do not affect the affinity of the 6xHis tag for the Ni-NTA resin, so purification of insoluble proteins is not a problem with affinity chromatography.
Spider silk proteins were purified either under native or denaturing conditions as described in Materials and Methods (Chapter 4). Table 2.1 gives an overview of all purifications carried out. To exemplify purification under native conditions Figure 4 shows the purification of the N- and C-terminal peptides of FLAG.

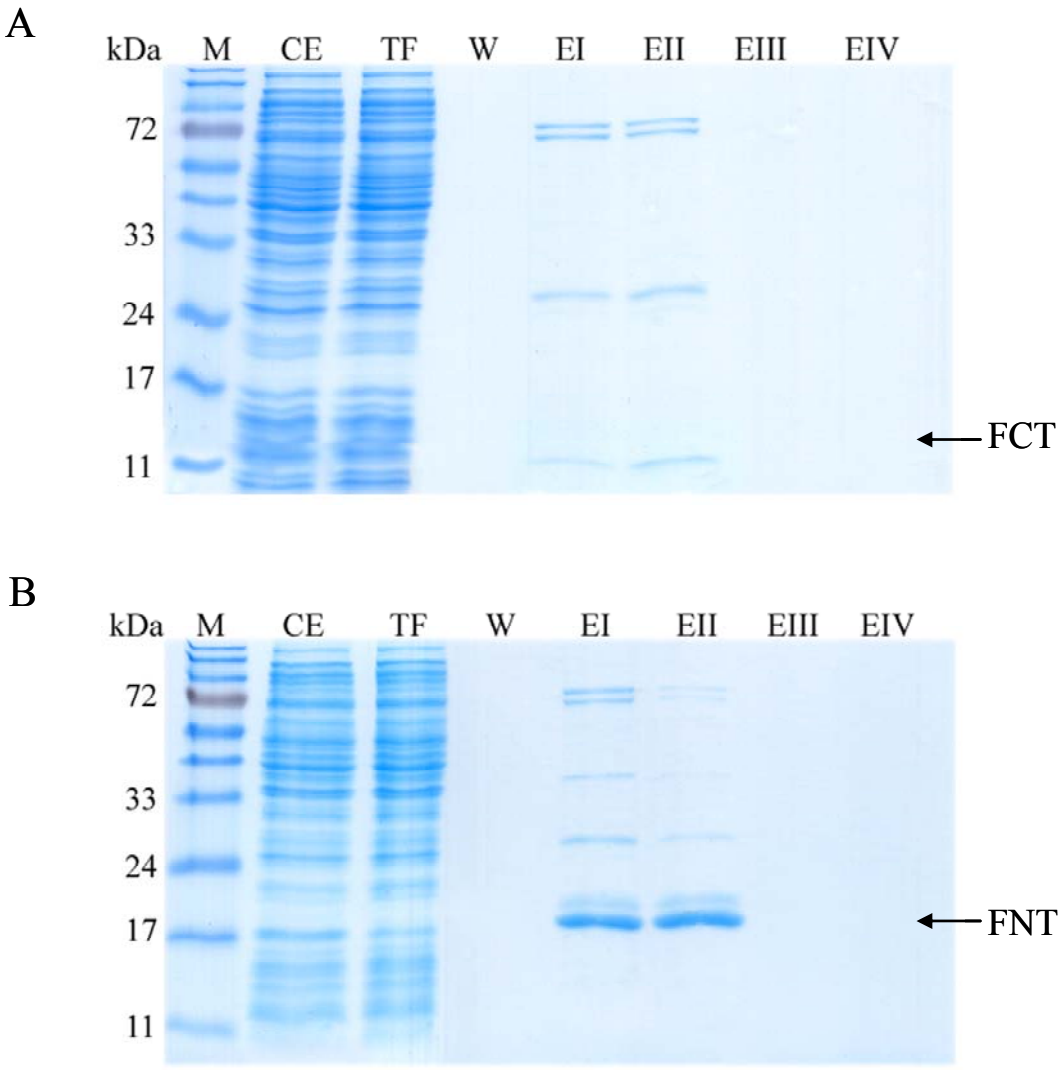


Figure 4. Purification of C-and N-terminal FLAG-derived spider silk peptides. (A) FCT. (B) FNT. Proteins were purified by native chromatography on a Ni^{2+} NTA column. Proteins were separated on SDS-PAGE and visualised by Coomassie staining. Lanes: CE: crude extract loaded on column; TF: flow through; W: wash with 30 mM imidazole; EI: elution with 100 mM imidazole; EII: elution with 200 mM imidazole; EIII and EIV: elutions with 250 mM imidazole.

The purification of spider silk proteins under denaturing conditions is demonstrated in Figure 5. Proteins from inclusion bodies were denatured by the presence of 6 M guanidine hydrochloride. Under these conditions target proteins were eluted already at lower imidazole concentration and were generally more pure compared to proteins purified under native conditions.

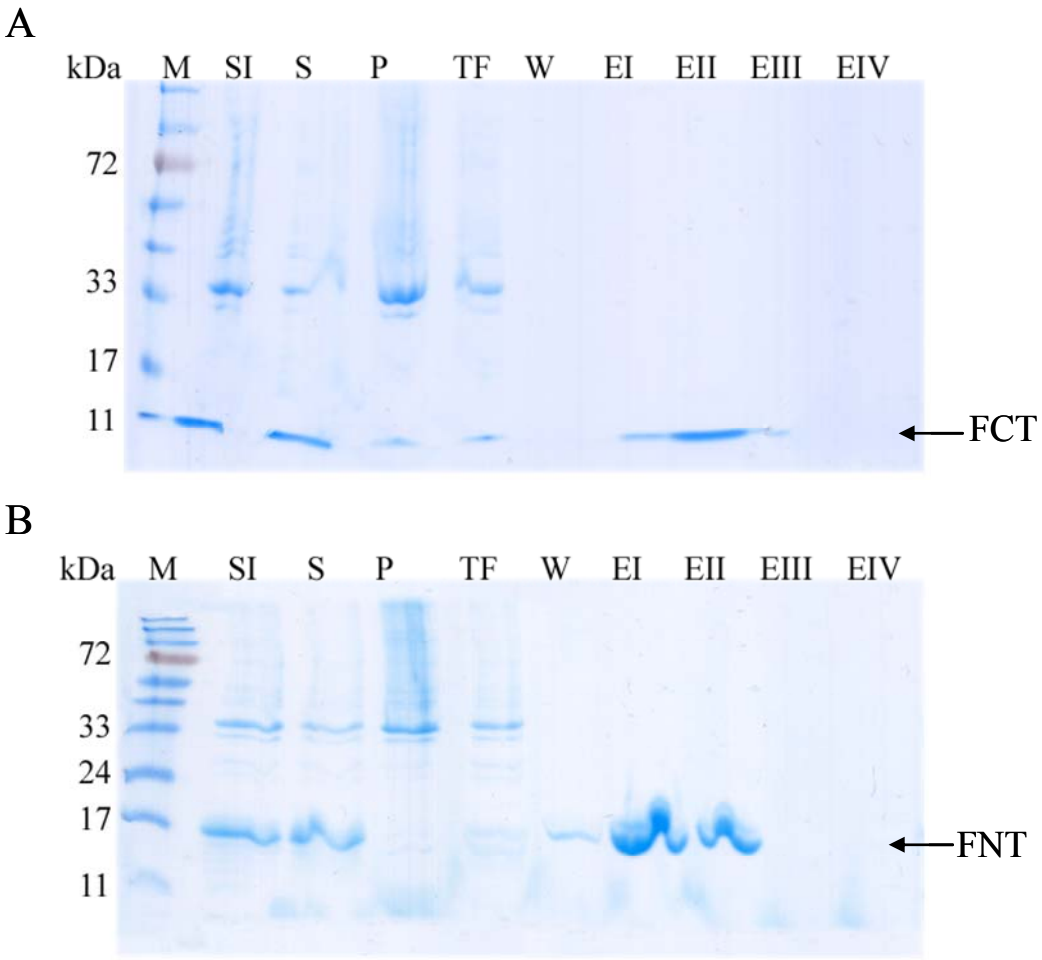


Figure 5. Purification of N- and C-terminal spider silk peptides from FLAG. Proteins were purified under denaturing conditions (6 M guanidine hydrochloride) on a Ni^{2+} NTA column. Proteins were separated on SDS-PAGE and visualised by Coomassie staining. (A) FCT. (B) FNT. Lanes: SI, solubilised inclusion bodies loaded on column; S, supernatant; P, pellet; TF, flow through; W, wash with 20 mM imidazole; EI, elution with 100 mM imidazole; EII, elution with 200 mM imidazole; EIII and EIV, elutions with 100 mM imidazole.

In the flow through the band of the expected size is absent.This shows that FCT and FNT bind to the Ni-NTA agarose column. In the washing fractions only traces of the recombinant protein are seen. Only the first two eluted fractions contain sufficient amount of protein, whereas in the last two fractions no recombinant protein could be detected. At 30 mM imidazole bound proteins are efficiently eluted (Figures 4, 5). As demonstrated in Figures 4 and 5 N- and C-terminal peptides of spider silk proteins were purified from the bacterial expression system pET23a, either as a soluble protein or in a denaturated form made soluble with 6 M guanidine hydrochloride. In Table 5, all these bacterial expression experiments are summarised.

Table 5. Concentration* of purified N- and C-termini of spider silk proteins.

Protein	**Purification conditions**	**Total protein yield*, (mg)**
FNT	native	2,5
FNT	denaturated	4,6
FNT-28	native	1,4
FNT-28,	denaturated	8,6
FCT	native	1,4
FCT	denaturated	22,7
MaSp1 CT	native	1,5
MaSp1 CT	denaturated	1,6
MaSp2 CT	native	1,4
MaSp2 CT	denaturated	8,9
FNT-29	native	7,7
FNT	native	1,1
FNT-29/28	native	6,5

*) The yield from a 500 ml expression culture of *Escherichia coli* BL21(DE3) cells

2.2 Generation of antibodies specifically binding spider silk non-repetitive peptides

2.2.1 Production of polyclonal antibodies in rabbits

This chapter describes the production of polyclonal antisera by immunisation of rabbits with antigens. Production of specific and high affine antisera mainly depends on the quality, purity, and the amount of available antigen. Such sera then are an essential base for the development of specific and sensitive immunological assays. For protein antigens, if possible, the material should be biochemically homogenous and, depending on the intended use, should be in either a native or denaturated conformation.
Antisera to be used for immunoblots are best made against denatured protein.
Production of polyclonal antisera takes less time and effort than production of monoclonal antibodies, requires relatively simple and readily available equipment, and produces antibodies that can be used for immunoprecipitation, immunoblotting and enzyme-linked immunosorbent assay (ELISA).

The choice of the animal for the production of antibodies depends upon the amount of antiserum desired, the evolutionary distance between the species from which the protein of interest has been derived and the species of the animal to be immunised, and prior experience with the immunogens.
Rabbits provide as much as 25 ml of serum from each bleed without significant harmful effects. The bacterial expressed antigens were applied either as soluble native proteins or as protein bands isolated from SDS-polyacrylamide gels.
Because protein purification frequently denatures molecules and synthetic peptides usually do not achieve native conformations, immunisation with synthetic peptides and gel-purified proteins generally has resulted in the production of antibodies that recognise the antigen in its denatured form. In several cases denatured proteins were solubilised in 6 M guanidine hydrochloride, dialysed and separated on a SDS polyacrylamide gel. Resulting bands were then sliced out and used for immunisation. In order to produce antibodies specific for the putative signal sequence a synthetic peptide, comprising the 29 N-terminal amino acid residues of FLAG has been synthesised and also fused to KLH- (Keyhole Limpet Hemocyanin, isolated from the hemolymph of the mollusc *Megathura crenulata*) or BSA (PINEDA, Berlin). The peptide-KLH fusion has been used for immunisation. All the immunisation experiments performed are summarised in Table 6. The immunisation schedule is described in Material and Methods (chapter 4).

Table 6. Production of antisera against spider silk variants/ peptides in rabbits.

Antigen	**Number of rabbits**	**Number of immunisations**	**Names of sera obtained from rabbits**
FNT, native	2	6	50, 51
FCT, denaturated	1	6	52
29aa FLAG N-terminal peptide* conjugated with KLH	2	3	66, 67

*) see Figure 1 B

2.2.2 Antibody detection and preparation

All sera have been tested by ELISA after the third immunisation. The results for FNT and FCT are shown in Figures 6, 7 and 8. At suitable dilutions specific binding against the corresponding antigens was detected in contrast to serum of non-immunised rabbits. The sera against the selected small peptide from the FLAG N-terminal non-repetitive part did neither bind to the peptide-KLH conjugate nor to FNT (data not shown). Furthermore, the sera No. 50, 51 and 52 were tested in Western blot according their binding to the corresponding antigens. Serum No. 50 bound to FNT and could detect at least 10 ng of antigen (Figure 9). The tests of serum No. 51 in Western blot resulted in very faint bands even at suitable serum dilutions (data not shown). Serum No. 52 did not show any reaction in Western blot (data not shown). Serum No. 50 was used successfully to detect

recombinant FLAG proteins with N-terminus in transgenic plants (Figure 26, 27, 28, 29). The serum No. 50 is an essential tool to characterise FLAG-transgenic plants without c-myc tag (see below).

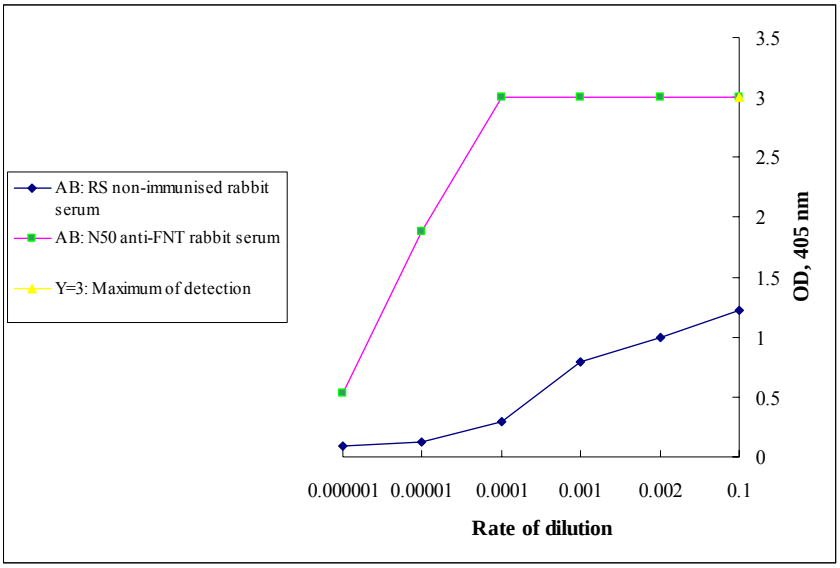


Figure 6. Detection of specific rabbit antibodies binding to the native FLAG non-repetitive N-terminal peptide. FNT was adsorbed to the plate, serum dilutions applied and binding detected by anti rabbit ALP and pNP phosphate.

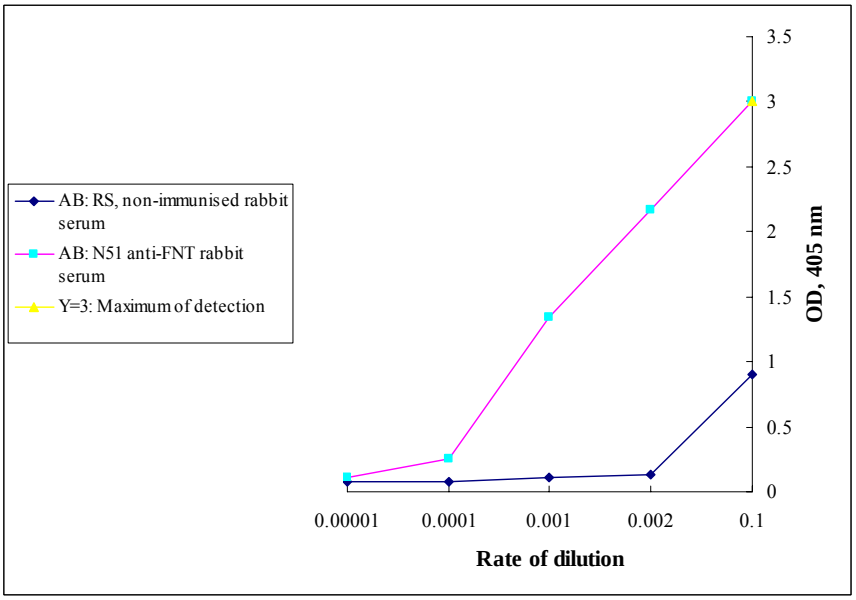


Figure 7. Detection of specific rabbit antibodies binding to native FLAG non-repetitive N-terminal peptide. FNT was adsorbed to the plate, serum dilutions applied and binding detected by anti rabbit ALP and pNP phosphate.

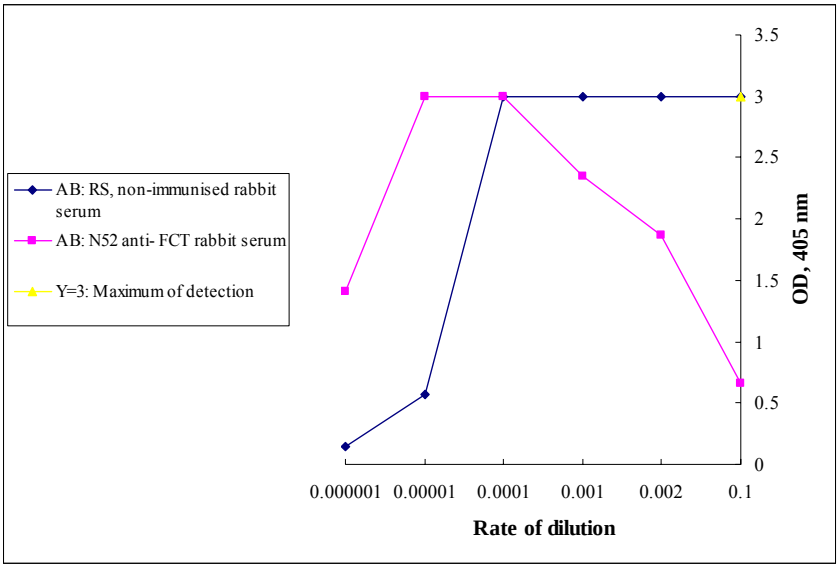


Figure 8. Detection of specific rabbit antibodies binding to the denatured FLAG non-repetitive C-terminal peptide. FCT in guanidine hydrochloride was diluted in PBS, pH 7.2 and adsorbed to the plate, serum dilutions applied and binding detected by anti rabbit ALP and pNP phosphate.

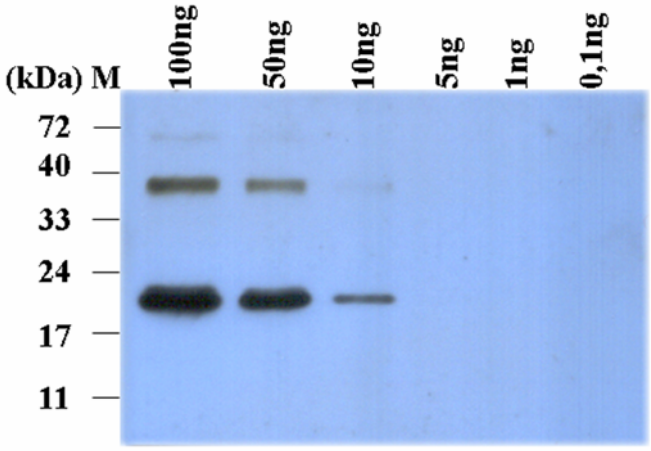


Figure 9. Detection of binding of rabbit anti FNT serum No. 50 to different amounts of FNT in Western blot. Antigens were separated by SDS-PAGE, blotted and detected by donkey anti rabbit POD conjugate and ECL.

2.2.3 Production of recombinant scFv antibodies *via* Phage Display

A powerful way to select specific recombinant antibodies is the so-called Phage Display technology. The display of peptides or proteins on phage surfaces has been described for the first time by G.P. Smith in 1985 (Smith 1985). In 1990 J. McCafferty and co-workers were the first to show the successful display of functional antibody fragments on phage

surfaces (McCafferty et al. 1990) and 1991 the antigen-driven selection of specific recombinant antibodies from Phage Display libraries was described (Clackson et al. 1991). Since this time, several other antibody formats as VH, VHH (single VH domains with full antigen binding properties and lacking any interface for VL binding) and Fab (heterodimeric antibody fragment) have been displayed and different types of libraries have been developed and used (for review see Conrad and Scheller 2005). For the screening experiments described here 2 single framework libraries (A and B) (de Wildt et al. 2000) were used. Immunoselections were performed as described by Gahrtz and Conrad (2007). After three rounds of selection single clones were selected and soluble recombinant antibodies were produced and tested for specific binding by ELISA. The determination of the sequence characterises the selected clones. These recombinant antibodies can be produced in bacterial systems as the pIT2 plasmid or in yeasts as *Pichia pastoris* and purified *via* the His tag using affinity chromatography.
The Phage Display experiments are described in Table 7. For the selection, native antigens FNT, denaturated antigens (C-terminal part of ADF-4(NR4), C-terminal part ADF-3 (NR3), CT from FLAG) or peptide conjugates (FNT-KLH, FNT-BSA) were used. ADF-3 and ADF-4 (fibroin from *Araneus diadematus*, Guerette et al. 1996), were kindly provided by Dr. Th. Scheibel (TU Munich). Specific soluble scFv/VH were isolated against the N-terminal part of FLAG (FNT), against the denaturated C-terminal part of FLAG (FCT) and against the C-terminal non-repetitive part of ADF-3

Table 7. Immunoselection of recombinant scFv antibodies by Phage Display.

Antigen	**Concentration used for coating**	**selected scFv/VH**
FNT, native	20 µg/ml	VH12,VH14,VH15,VH16,VH18, ScFv21,ScFv22,VH23
FCT, denaturated	20 µg/ml	ScFv1, VH2, VH3,VH5,VH6,VH7,VH8
CT of ADF-4 (NR4), denaturated	10 µg/ml	none
CT of ADF-3 (NR3), denaturated	10 µg/ml	VH22, VH5, VH10, VH19, VH20, VH21, VH3, ScFv23
FNT peptide* KLH-conjugate (FNT-KLH)	5 µg/ml	none
FNT peptide BSA-conjugate (FNT-BSA)	5 µg/ml	none

*) FNT peptide is a 29 amino acid sequence from amino-termini of FLAG N-terminus (see Figure 1B)

The clones selected after immunoselection with FNT, FCT and CT of ADF-3 have been produced as soluble recombinant antibodies and tested by an indirect ELISA. The results are shown in Figures 10, 11 and 12, respectively. Specific binding to the corresponding antigens could be verified.

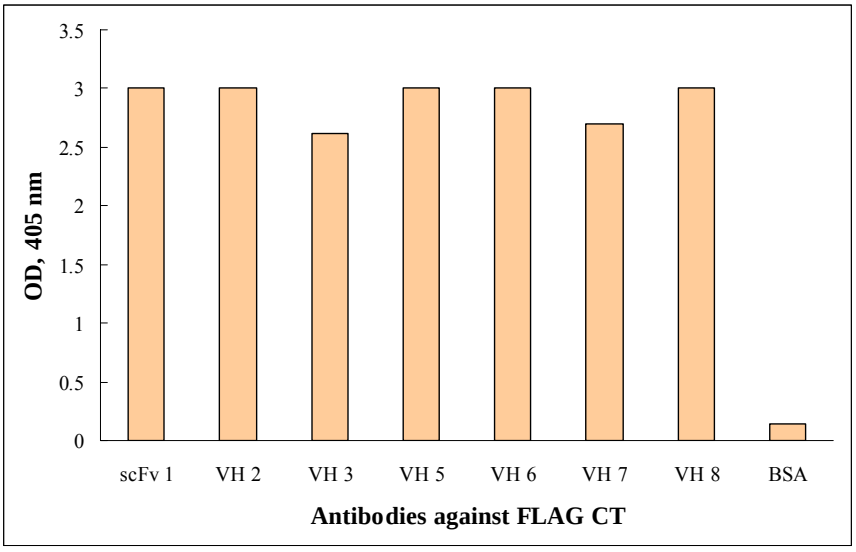


Figure 10. Specific binding of anti FLAG CT scFv/VH to the antigen FCT in indirect ELISA. Antigen concentration was 20 µg/mL for FCT or 10 µg/mL for BSA (control). Binding of phagemid supernatant was detected *via* anti c-myc antibody, rabbit anti mouse-IgG conjugated ALP and pNPP.

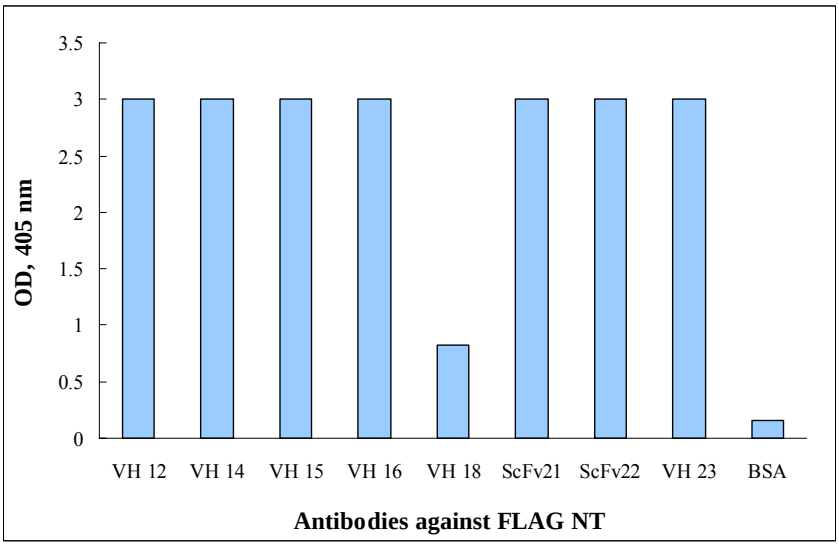


Figure 11. Specific binding of anti FLAG NT scFv/VH to the antigen FNT in indirect ELISA. Antigen concentration 20 µg/mL for FNT or 10 µg/mL for BSA (control). Binding of phagemid supernatant was detected *via* anti c-myc antibody, rabbit anti mouse-IgG conjugated ALP and pNPP.

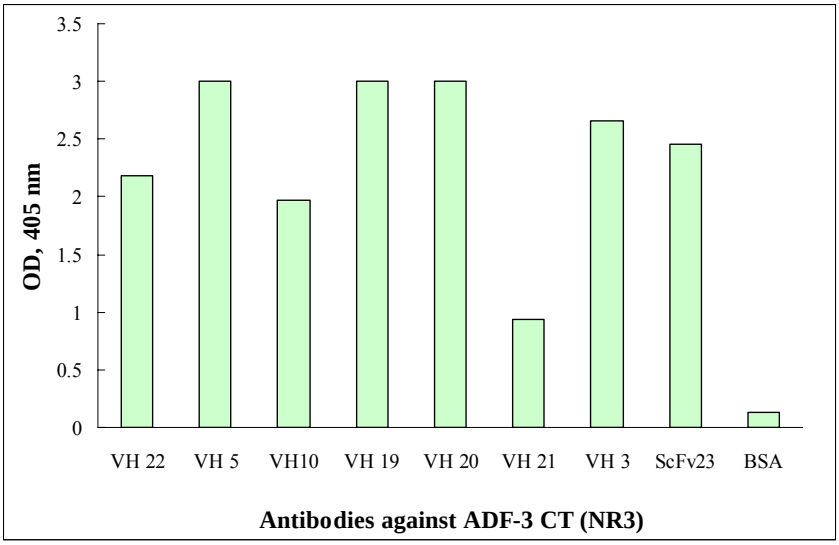


Figure 12. Specific binding of anti ADF-3 CT (NR3) scFv/VH to antigen ADF-3 CT (NR3) in indirect ELISA. Antigen concentration 20 µg/mL for ADF-3 CT (NR3) or 10 µg/mL for BSA (control). Binding of phagemid supernatant was detected *via* anti c-myc antibody, rabbit anti mouse-IgG conjugated ALP and pNPP.

The cross-reactivity of selected clone has been further tested. The anti FLAG antibodies have been cross-tested with each other. No cross-reactivity was detected (data not shown). Because a sequence homology of the C-terminus of ADF-3 to MaSp1 and MaSp2 is known (Figure 1D), the anti ADF-3 antibodies have been tested against the C-terminus of MaSp1, against the C-terminus of MaSp2 as well as against the C- and N-termini of FLAG to detect possible cross-reactivity. As shown in Figure 13, the recombinant antibodies tested here (anti ADF-3 (NR3) VH3, VH20, VH22 and scFv23) cross-react with the conserved regions of the C-terminus of MaSp1.

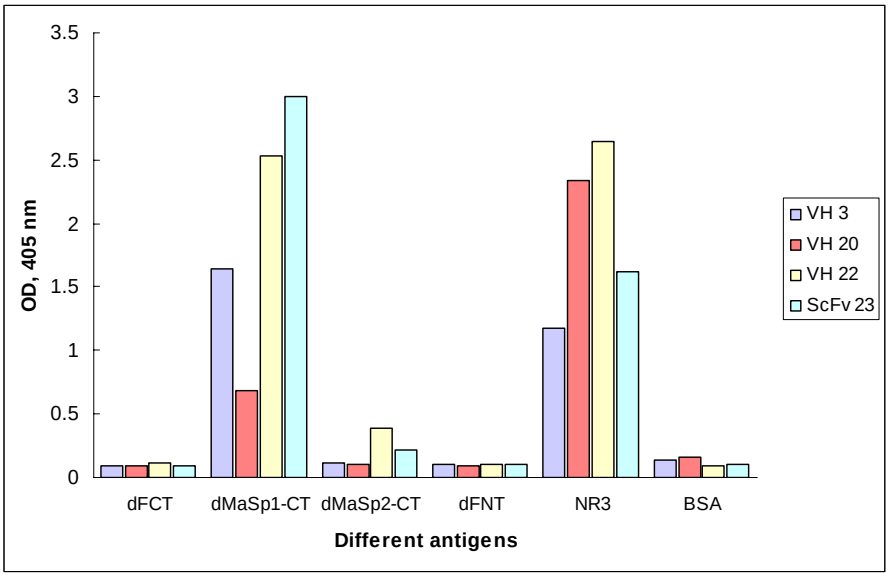


Figure 13. Specific binding of anti ADF-3 CT (NR3) VH/ScFv to antigens ADF-3 CT and dMaSp1 CT in indirect ELISA. Antigen concentration 20 µg/mL for ADF-3 CT (NR3) or 10 µg/mL for BSA (control). Binding of phagemid supernatant was detected *via* anti c-myc antibody, rabbit anti mouse-IgG conjugated ALP and pNPP.

Recombinant antibodies have all been tested by Western blot analysis. Only one (anti FLAG NT VH15) was specifically binding its antigen in Western blot (Figure 14). It is a general phenomenon that even scFv/VH with high affinity in ELISA do not bind to antigens adsorbed to nitrocellulose membranes. The epitope of the antibody anti FLAG NT VH15 is located between the amino acids 30 and 98 within the core region.

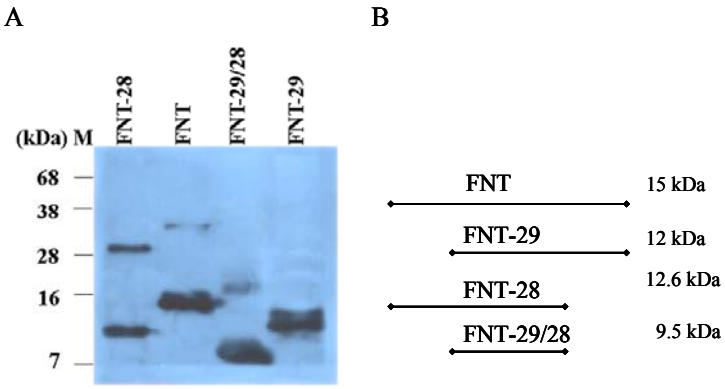


Figure 14. Characterisation of anti FNT VH15 antibody in Western blot. (A): detection of antigen FNT variants by VH 15, monoclonal mouse anti c-myc tag antibodies, sheep anti mouse POD conjugate and ECL. (B): Antigen fragments (FNT: full size of non-repetitive FLAG N-terminus (1-126aa); FNT-29: truncated FLAG N-terminus from amino terminus (30-126aa); FNT-28: truncated FNT from carboxy terminus (1-98aa); FNT-29/28: truncated FNT from both amino- and carboxy- parts (30-98aa). Sizes of antigens indicated in kDa.

Selected scFv/VH has been sequenced. The results are summarised in Figure 15.

A

```
                        Start of VH     FR-H1           CDR-H1    FR-H2
FNT-ScFv21-1pIT2  (1)  MAEVQLLESGGGLVQPGGSLRLSCAASGFTFSGYGMHWVRQAPGKGLEWV
NR3-ScFv23- pIT1  (1)  MAEVQLLESGGGLVQPGGSLRLSCAASGFTFSNYSMRWVRQAPGKGLEWV

                           CDR-H2                   FR-H3
FNTScFv21-1pIT2  (51)  SAISGSGGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK
NR3-ScFv23-pIT1  (51)  SAISGSGGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK

                       CDR-H3     FR-H4     Linker    Start of VL  FR-L1
FNTScFv21-1pIT2  (101)PRTTFDYWGQGTLVTVSSGGGGSGGGGSGGGGSTDIQMTQSPSSLSASVG
NR3-ScFv23-pIT1  (101)NLIVFDYWGQGTLVTVSSGGGGSGGGGSGGGGSTDIQMTQSPSSLSASVG

                                  CDR-L1     FR-L2    CDR-L2    FR-L3
FNTScFv21-1pIT2  (151)DRVTITCRASQSILKYLTWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSG
NR3-ScFv23-pIT1  (151)DRVTITCRASQSIRDELKWYQQKPEKAPKLLIYAASSLQSGVPSRFSGSG

                                                CDR-L3     FR-L4
FNTScFv21-1pIT2  (201)SGTDFTLTISSLQPEDFATYYCQQSYSTPPTFGQGTKVEIKRAAA
NR3-ScFv23-pIT1  (201)SGTDFTLTISSLQPEDFATYYCQQSYSTPHTFGQGTKVEIKRAAA
```

B

```
                        Start of VH     FR-H1           CDR-H1    FR-H2
FCT-VH2-2 pIT2    (1)  MAEVQLLESGGGLVQPGGSLRLSCAASGFTFSKYFMAWVRQAPGKGLEWV
FNT-VH15 pIT1     (1)  MAEVQLLESGGGLVQPGGSLRLSCAASGFTFSAYNMTWVRQAPGKGLELV
NR3-VH22 pIT1     (1)  MAEVQLLESGGGLV-PGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWV

                         CDR-H2                        FR-H3
FCT-VH2-2 pIT2   (51)  SAISGSGGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAE
FNT-VH15  pIT1   (51)  SAISGSGGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK
NR3-VH22  pIT1   (50)  SRIFQWGKSTDIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKP

                       CDR-H3     FR-H4     Linker
FCT-VH2-2 pIT2  (101)  WVTTFDYWGQGTLVTVSSGGGGSGGGGSGGGGSTDIQMTQAAA
FNT-VH15  pIT1  (101)  TVRAFDYWGQGTLVTVSSGGGGSGGGGSGGGGSTDIQMTQAAA
NR3-VH22  pIT1  (100)  GKAPKLLIYRASSLQSGVPSRFSGSGSGTDFTLTISRLQPEDF
```

Figure 15. Alignment of amino acid sequences of (A): ScFv; (B): VH selected from phage displayed library against FNT: VH 15, ScFv 21-1; anti FCT: VH2; anti NR3: ScFv 23, VH 22. The CDR regions are highlighted in colors: red show identical, whereas blue, orange and green differences. The linker is underlined. VH and VL are heavy chain domain and light chain domain, respectively. FR-H1-4 and FR-L1-4 are framework regions of the variable domains.

As expected for clones from single framework libraries they differ only in the CDR regions. ScFv clone 21-1, VH clone 2, VH clone 22 used in further cloning and expression experiments in pIT2 and yeast vectors (see Figure 42 and 43 in Appendix). ScFv 21-1 and VH 2 were remained stably active, whereas VH 22 lost activity during expression in yeast. In VH22 a stop codon was identified but nevertheless the antibody fragment was expressed in *Escherichia coli*. This "overreading" of stop codons in HB2151 was observed several times already with scFv/VH from the libraries used here.

In case of further use for yeast expression, the stop codons have to be removed by *in vitro* mutagenesis.
Recombinant antibodies have been produced in bacteria and purified by metal affinity chromatography using the His tag. In selected cases, yeast expression experiments have been performed. To this aim, different scFv/VH coding genes (anti FNT VH 15, anti FNT ScFv 21, anti FCT VH 2) were cloned into the yeast expression vector pPICZ alpha B and transformed into *Pichia pastoris* by electroporation. Expression clones were selected by ELISA as it is exemplified for anti FLAG NT VH15 in Figure 16. The suitable dilution rate of expressing clones by ELISA is also demonstrated for anti FLAG NT VH15 (Figure 17).
Altogether, a panel of recombinant antibodies has been isolated and basically characterised. Together with the polyclonal antisera it opens the way to establish detection assays for monitoring spider silk-transgene expression even in cases were artificial tags have been avoided. Furthermore, the antibody collection could also be used for the analysis of spider silk non-repetitive N-and C-termini in spiders.

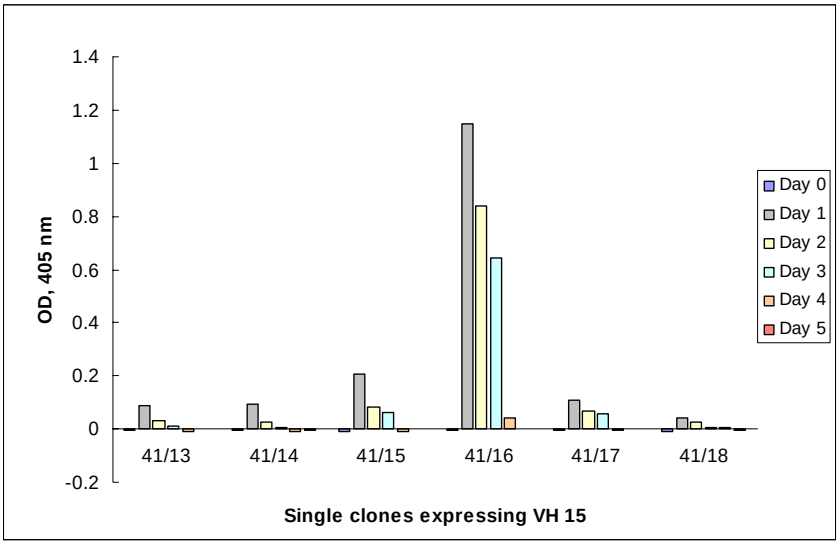


Figure 16. Selection of yeast expression clones by ELISA. Yeast clones expressing anti FLAG NT VH15 at different time points were tested against FNT. Binding is detected by anti c-myc antibody, rabbit anti mouse IgG conjugated ALP and pNPP.

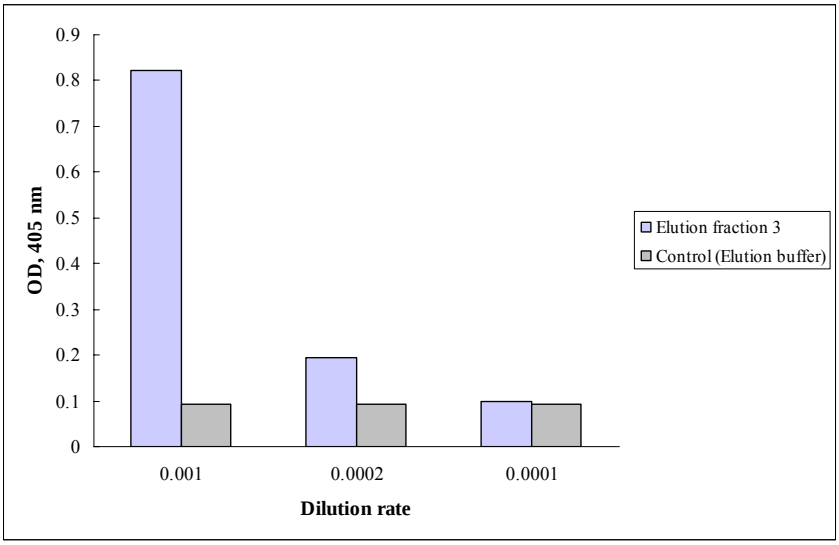


Figure 17. Testing of different dilution rates of yeast-expressed and purified anti FLAG VH 15 antibody by ELISA. Anti FLAG NT VH15 at different dilution rates were tested against FNT. Binding is detected by anti c-myc antibody, rabbit anti mouse IgG conjugated ALP and pNPP.

2.3 Production of spider silk and spider silk-ELP fusion proteins in plants

At the time when this work has been started, only MaSp1-derivatives SO1 (Scheller et al. 2001) have been produced in plants. Later, SO1 has been fused to elastin-like peptides to enhance expression and to develop a reliable purification system of practical use (Scheller et al. 2004). Both tobacco and potato (*Solanum tuberosum*, cv. Solara) have been transformed to produce spider silk derivatives (Zambryski et al. 1983). For the expression of spider silk proteins in plants the transgene can be reduced to only the repetitive parts, as it has been done earlier (Scheller et al. 2001, Menassa et al. 2004). Here MaSp1 and MaSp2 repetitive sequences were used to generate ELP fusions. In contrast to that spider silk sequences comprising the non-repetitive N- and C-terminal parts haven't been expressed and characterised up to now. Therefore, a spider silk sequence with both natural non-repetitive N-and C-terminal sequences should be designed and cloned into a plant expression vector. We have chosen the *FLAG* gene from *Nephila* for this purpose, because, at that time when this work has been started, this was the only available spider silk gene with known N-and C-terminal sequences. In addition to tobacco starch potatos, bred for technical use, were transformed to evaluate potato as a production system.

2.3.1 Design of the expression vectors and transient expression of recombinant FLAG proteins

Generally, gene fusions and other manipulations are done in the vector pGEM-T (www.promega.com), because this plasmid contains a rather extended polylinker and is easy to handle in terms of DNA preparation and sequencing. Besides that, the pGEM-T

vector cloning system allows direct cloning of PCR products using the T/A cloning strategy. Expression cassettes are then designed in the vector pRTRA 7/3 (Artsaenko et al. 1995). In this cassette, the legumin B4 signal peptide coding sequence provides translation at the ER associated ribosomes and entry in the ER. The KDEL coding sequence causes ER retention and the c-myc tag serves as suitable antigen for detection using Western blot and ELISA. From here, HindIII fragments can be easily transferred into the shuttle vector pCB301 Kan (Scheller et al. 2004). The expression vectors for the production of MaSpI-100xELP, MaSp2-100xELP and SO1-100xELP were kindly provided by Dr. Jürgen Scheller (Figure 18). The design of the variants of *FLAG* expression vectors is described in the following chapter.

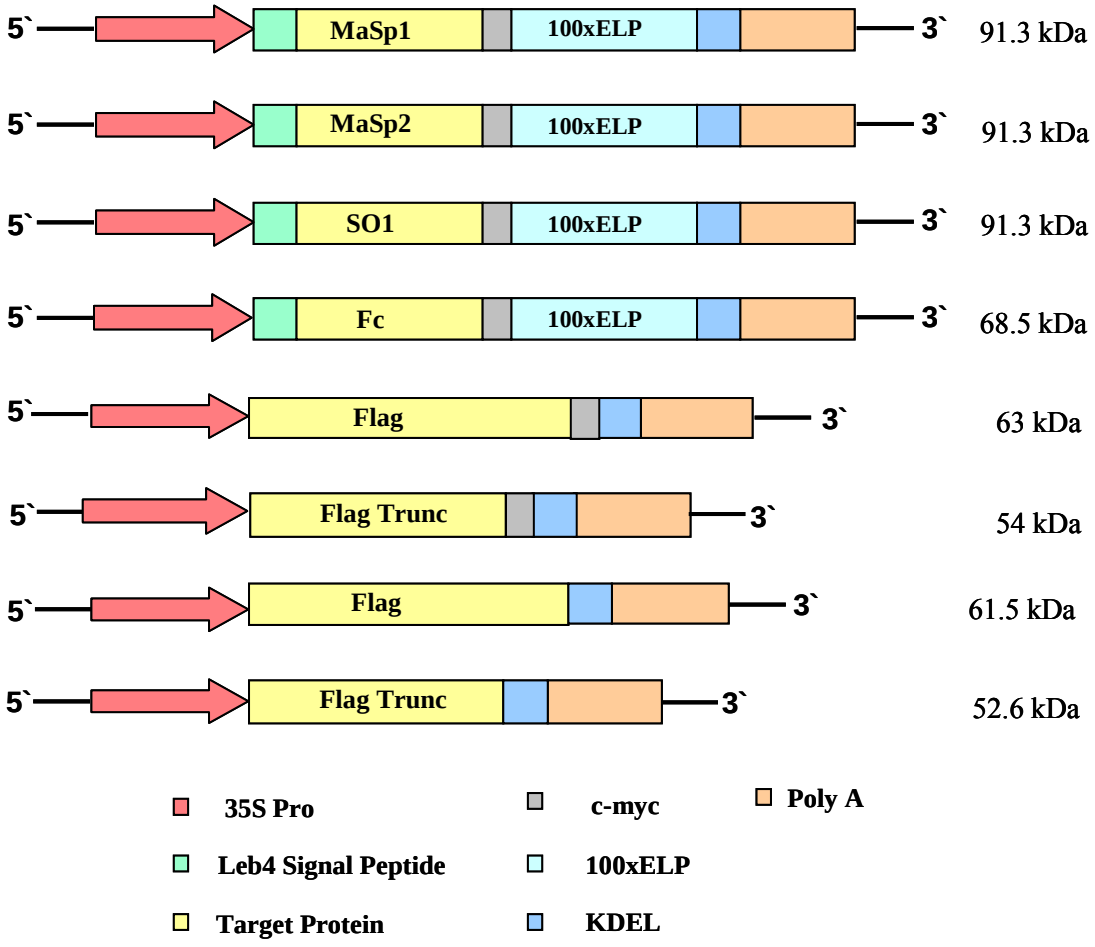


Figure 18. Schematic representations of plant expression cassettes for ubiquitous production of different spider silk and spider silk like genes.

2.3.1.1 Cloning of the FLAG protein encoding gene into the pGEM-T vector

The cloning experiments are performed in several steps. At first, the 5´ non-repetitive part was amplified by PCR, then the 3´non-repetitive part was amplified and at last, the

repetitive part was inserted resulting in a construct where non-repetitve parts and one complete ensemble repeat are now combined in the vector pGEM-T. The *FLAG* N-terminal gene fragment was amplified by PCR using the FOR_Fl_N-term and BACK_Fl_N-term primers containing SmaI and MunI sites, respectively (Figure 20A). The 943 bp PCR product was ligated into pGEM-T by T/A cloning. The resulting vector was called pGEM-T Flag NT-RP. This fragment of the *FLAG* DNA contains the N-terminus of FLAG and a 5`fragment of the repetitive part of FLAG (Figure 20B). Similarly, the *FLAG* C-terminal gene fragment was amplified by PCR using the FOR_Fl_C-term and BACK_Fl_C-term primers containing MunI and NaeI sites (Figure 20C). The PCR product (743 bp) was also ligated using T/A cloning into pGEM-T. The resulting vector was called pGEM-T Flag CT (Figure 20D). The plasmid pCYH1 was digested by EcoNI and MunI restriction enzymes. A released 843 bp fragment of the *FLAG* gene encoding the repetitive part of the protein was eluted and inserted into the vector pGEM-T Flag NT-RP, cut by EcoNI and MunI. The resulting plasmid pGEM-T Flag NT-RP was 4757 bp in length (Figure 20B). The C-terminal fragment of *FLAG* was cut out from the pGEM-T Flag CT by MunI and PstI sites and ligated into the vector pGEM-T Flag NT-RP cut by MunI and PstI. The resulting plasmid was 5477 bp in length and named pGEM-T NT-RP-CT and coding for the N-terminus, a repetitive part and the C-terminus of FLAG, which codes for a recombinant protein of 747 amino acids (Figures 19 and 20E).

```
  1    MGKGRHDTKA KAKAMQVALA SSIAELVIAE SSGGDVQRKT NVISNALRNA
 51    LMSTTGSPNE EFVHEVQDLI QMLSQEQINE VDTSGPGQYY RSSSSGGGGG
101    GQGGPVVTET LTVTVGGSGG GQPSGAGPSG TGGYAPTGYA PSGSGAGGVR
151    PSASGPSGSG PSGGSRPSSS GPSGTRPSPN GASGSSPGGI APGGSNSGGA
201    GVSGATGGPA SSGSYGPGST GGTYGPSGGS EPFGPGVAGG PYSPGGAGPG
251    GAGGAYGPGG VGTGGAGPGG YGPGGAGPGG YGPGGAGPGG YGPGGAGPGG
301    YGPGGAGPGG YGPGGAGPGG YGPGGAGPGG YGPGGTGPGG YGPGGTGPGG
351    VGPGGAGPGG YGPGGAGPGG AGPGGAGPGG AGPGGAGPGG AGPGGYGPGG
401    SGPGGAGPSG AGLGGAGPGG AGLGGAGPGG AGTSGAGPGG AGPGGAGQGD
451    AGPGGAGRGG AGRGGVGRGG AGRGGAGRGG ARGAGGAGGA GGAGGSGGTT
501    IVEDLDITID GADGPITISE ELTISGAGGS GPGGAGPGGV GPGGSGPGGV
551    GPGVSGPGGV GPGGSGPGGV GSGGSGPGGV GPGGYGPGGS GSGGVGPGGY
601    GPGGSGGFYG PGGSEGPYGP SGTYGSGGGY GPGGAGGPYG PGSPGGAYGP
651    GSPGGAYYPS SRVPDMVNGI MSAMQGSGFN YQMFGNMLSQ YSSGSGTCNP
701    NNVNVLMDAL LAALHCLSNH GSSSFAPSPT PAAMSAYSNS VGRMFAY
```

Figure 19. Deduced amino acid sequence of recombinant FLAG with its non-repetitive N-terminus and C-terminus highlighted in red and the repetitive part represented by one ensemble repeat as encoded by the respective open reading frame in the plasmid pGEM T NT-RP-CT

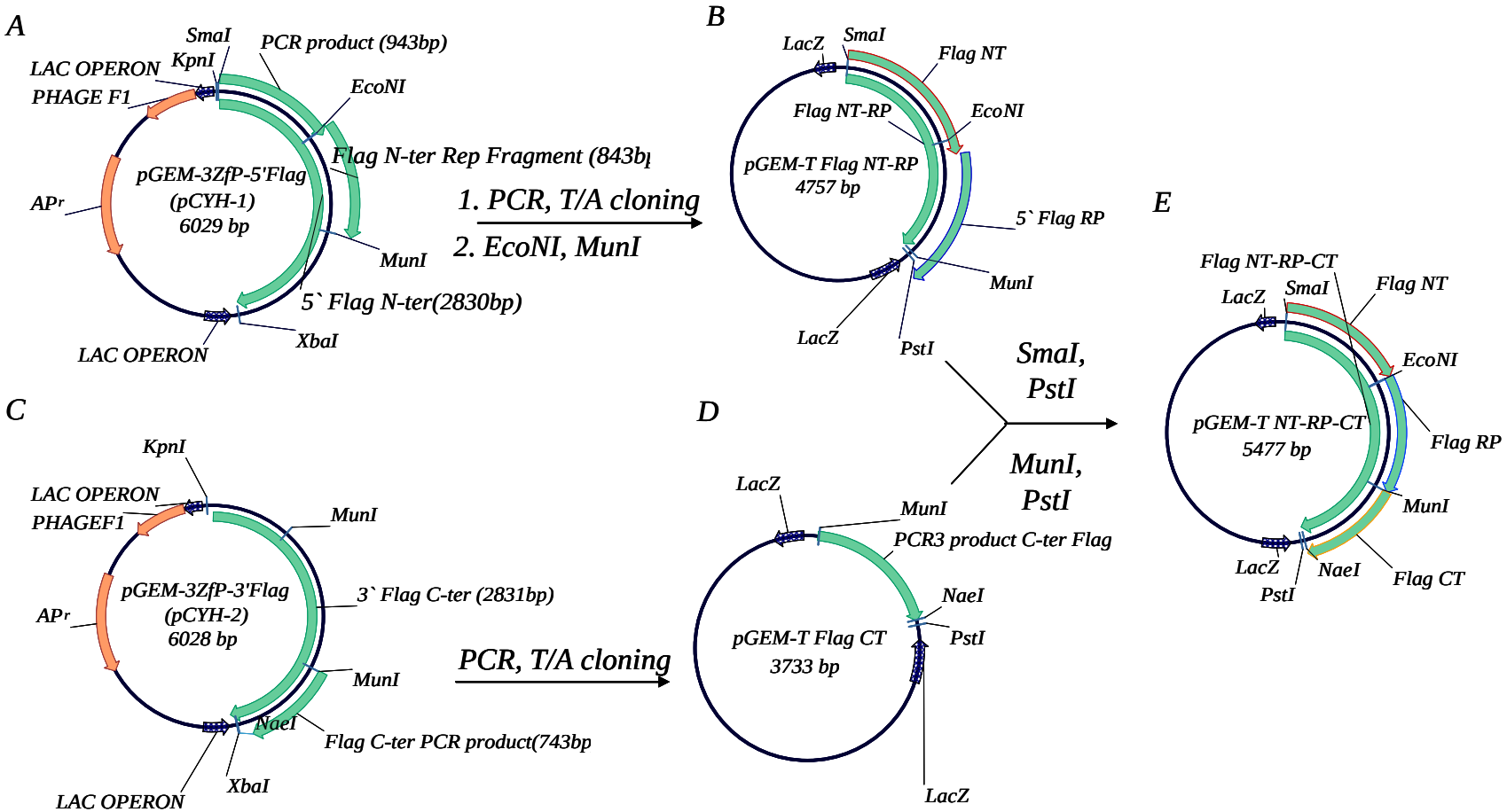


Figure 20. Construction of the recombinant *FLAG* gene containing the non-repetitive N-terminus, the non-repetitive C-terminus and one ensemble repeat. (A): pCYH1 plasmid containing cDNA sequences corresponding to FLAG's 5` N-terminal part; (B): pGEM-T vector containing non-repetitive N-terminus and 5`repetitive part of *FLAG* gene; (C): pCYH2 plasmid containing cDNA sequences corresponding to FLAG`s 3` carboxy terminus. (D): pGEM-T vector containing non-repetitive C-terminus of *FLAG*. (E): pGEM-T vector containing non-repetitive N-terminus, C-terminus and one ensemble repeat.

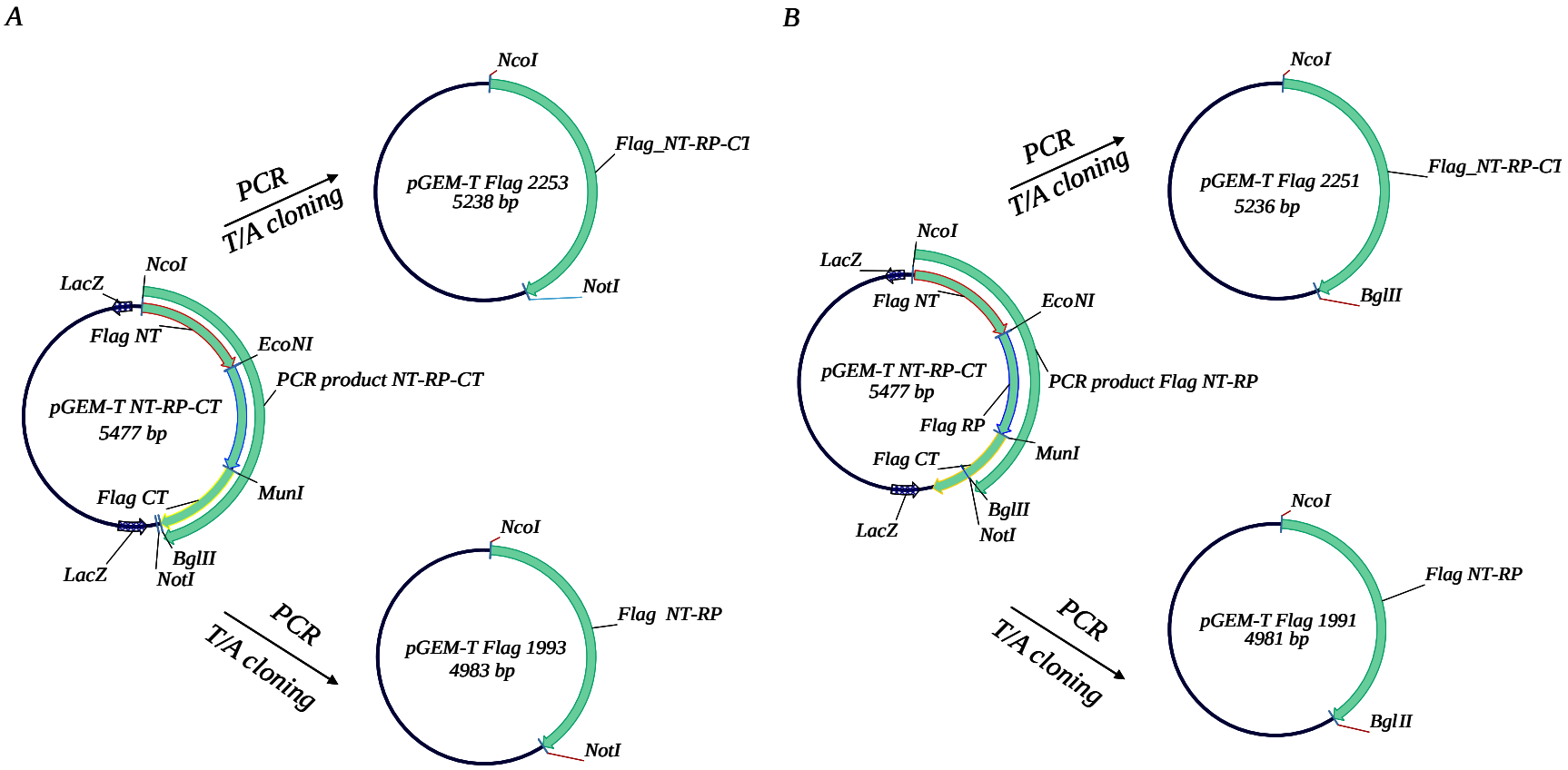


Figure 21. Scheme representing introduction of approapriate restriction enzyme sites for subsequent cloning of *FLAG* gene-variants into pRTRA 7/3 vector. (A): PCR products with NcoI and NotI sites; (B): PCR products with NcoI and BglII sites.

In order to introduce the recombinant *FLAG* gene in a plant expression cassette appropriate restriction sites for the cloning into the plasmid pRTRA 7/3 had to be introduced at the ends of the *FLAG* gene. Therefore, the four different PCR products were synthesised using the primers FOR Flag_N-ter and the four reverse primers BACK Flag_c-myc C-ter, BACK Flag_Trunc c-myc_C-ter, BACK _Flag_C-ter, BACK Flag_Trunc_C-ter (see Materials and Methods) and were cloned into pGEM-T. The vector pGEM-T NT-RP-CT Flag was used as PCR-template. The resulting plasmids were named pGEM-T Flag2253, pGEM-T Flag1993, pGEM-T Flag2251, pGEM-T Flag1991 (Figure 21). The sequences of these constructs were verified by sequencing with the primers SP6 and universal M13 forward (see Materials and Methods). The deduced amino acid sequence was found to be as predicted from the originating sequences from *Nephila clavipes* encoding for a FLAG protein comprising a complete ensemble repeat including the non-repetitive N- and C-terminal sequences.

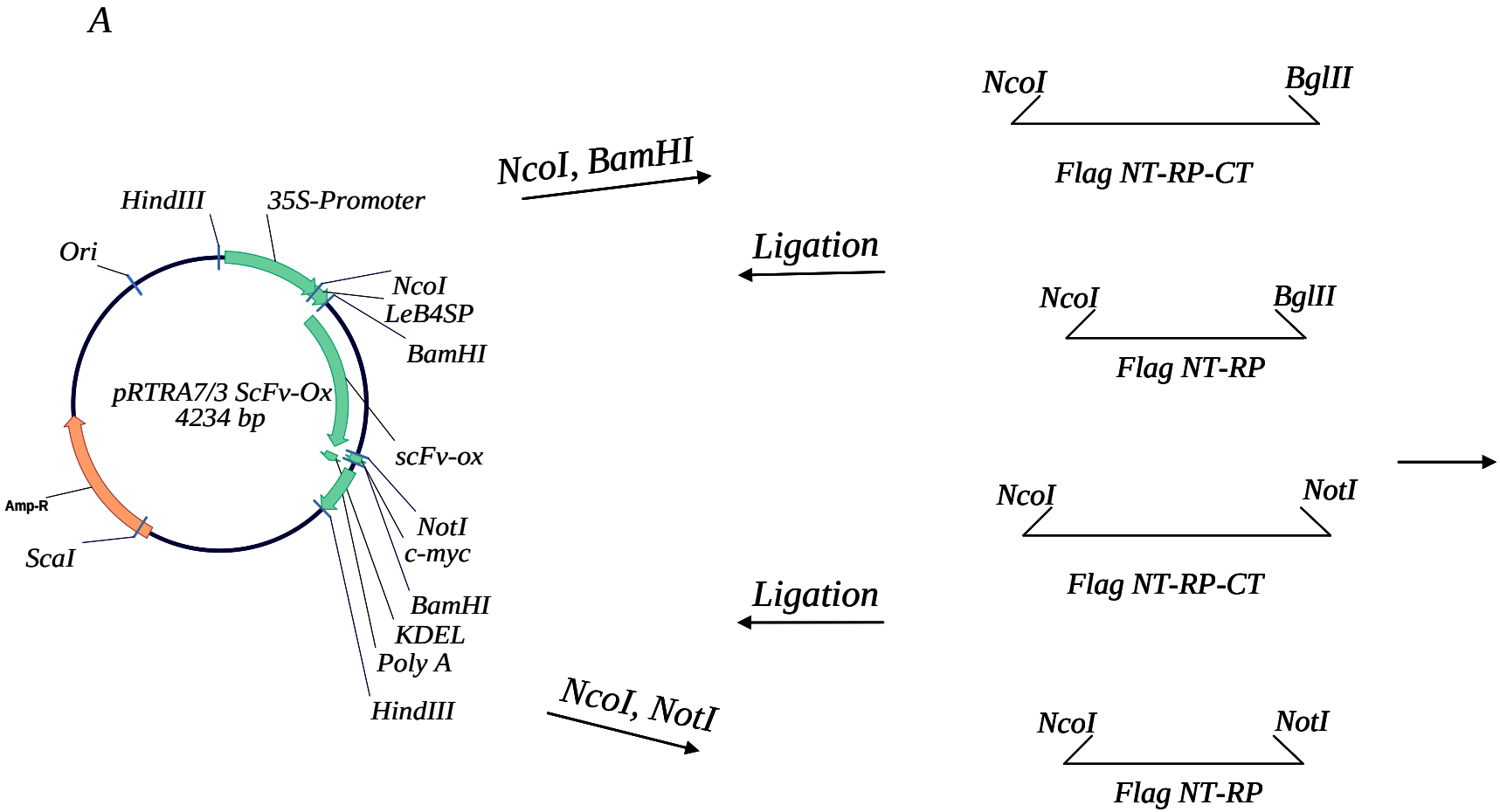


Figure 22A. Cloning of *FLAG* gene variants into pRTRA 7/3 vector in frame with KDEL and c-myc sequences.

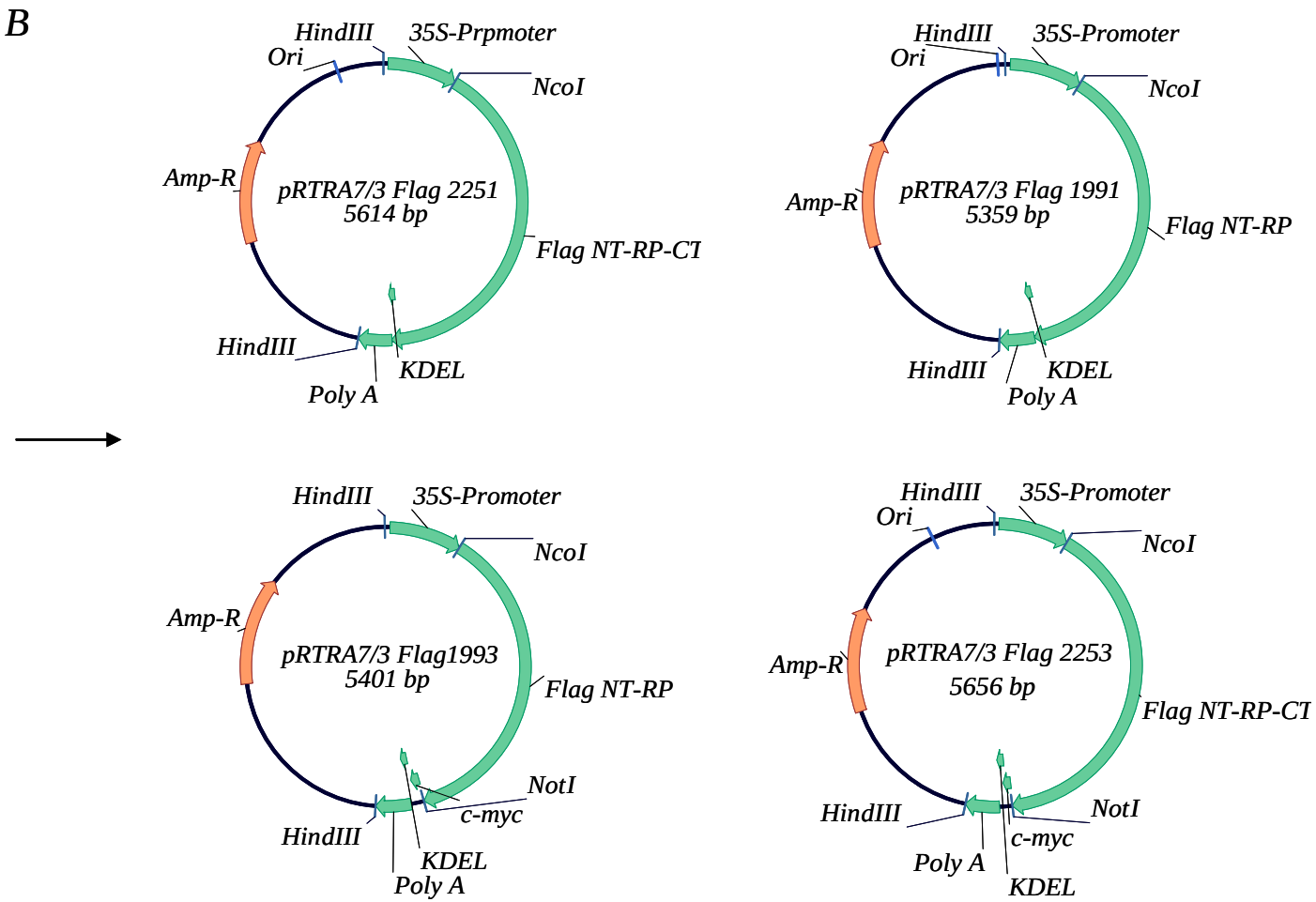


Figure 22B. Construction of pRTRA 7/3 vector containing variants of *FLAG* gene with and without c-myc tag.

2.3.1.2 Cloning of the *FLAG* gene and of *FLAG* gene fragments into the pRTRA 7/3 vector and transient expression in tobacco protoplasts

The basic pRTRA 7/3 vector contains an anti ABA scFv gene with the legumin signal sequence, the c-myc tag and the KDEL encoding sequence fused to the 3`-end of the tag (Artsaenko et al. 1995). Transcription is driven by the CaMV 35S promoter (Toepfer et al. 1993). The tetrapeptide KDEL has been shown to be sufficient for the retention of a number of proteins in the ER (Inohara et al. 1989) including spider silk proteins (Scheller et al. 2001). Moreover, increased stability and accumulation of proteins containing carboxy-terminal KDEL in transgenic plants has been reported (Wandelt et al. 1992).The legumin signal sequence is responsible for the transport of proteins through the ER membrane (Saalbach et al. 1991).The c-myc tag is used to enable rapid and sensitive detection, purification and quantification of the recombinant spider silk proteins by immunological methods. This tag is widely applied in protein expression studies including plants as well (Hoogenboom et al. 1991, Artsaenko et al. 1994, Fiedler and Conrad 1995, Scheller et al. 2001). This peptide contains 11 amino acids fused to the C-terminus of the recombinant proteins and is recognised by the 9E10 antibody (Munroe and Pelham 1986).
Recently, Foo and co-workers described detailed charge distribution in silk proteins and considered the predicted pI's of silk proteins and how pH should influence the overall charge on these proteins during *in vivo* processing. We performed theoretical calculation of the pI of the hypothetical FLAG protein variants by using the BioPlot application of vector NTI 6.0 program to find out if the addition of c-myc tag sequences to the C-terminus of FLAG protein will change the total net charge of protein (Table 8).

Table 8. Isoelectric point and charge determined for FLAG protein molecules without and with c-myc-tag

Translated sequence	Length, aa	Isoelectric point (pI)	Net charge at pH7
FLAG with c-myc tag*	734**	4.73	-9.98
FLAG without c-myc tag*	720**	5.05	-5.99

* see chapter 2.1
** predicted signal peptide is not included

The net charge at pH7 and the isoelectric point of the FLAG protein change remarkably, when the c-myc tag is added, as shown in Table 8. Therefore, we cloned a *FLAG* gene without fusion of any tag into the pRTRA 7/3 vector. For immunodetection of such proteins, which do not contain any tag we raised antibodies against non-repetitive parts of FLAG protein using powerful methods of recombinant antibody engineering (see chapter

Materials and Methods). In addition, the C-terminal non-repetitive parts shown to be *bona fide* parts of the mature spidroins and might serve to cross-link the spider silk fibroins by the formation of disulphide bridges (Sponner et al. 2004). Constructs with and without C-terminal non-repetitive sequences were designed to investigate the influence of these peptide sequences on protein properties after plant expression. All the different variants are shown at Figure 18.

The FLAG protein from *Nephila clavipes* also contains a non-repetitive N-terminal peptide sequence. A possible signal sequence of 29 amino acids is predicted in this peptide (Hayashi et al. 1998). Homologous sequences even from other spider silk proteins at the N-terminus show signal peptide probability scores above threshold (Rising et al. 2006). We avoided the Leb4 signal peptide from the pRTRA 7/3 vector by using NcoI cleavage to check in transgenic plants, if this 29 amino acid sequence functions as signal peptide. In order to avoid the c-myc tag at the 3` end, the vector pRTRA 7/3 was cut by BamHI. In case when we used c-myc tag as fusion to the C-terminus of FLAG, the pRTRA 7/3 vector was digested by NotI. In order to place the designed *FLAG* gene into the pRTRA 7/3 vector, pGEM-T Flag2253 and pGEM-T Flag1993 were cut by NcoI and NotI and excised fragments were inserted into the pRTRA 7/3 vector predigested with NcoI and NotI (Figure 22A). Similarly, pGEM-T Fag2251 and pGEM-T Flag1991 were cut by NcoI and BglII and yielded fragments were ligated into the pRTRA 7/3 vector, which was cut by NcoI and BamHI restriction enzymes. Sticky ends of BamHI and BglII are compatible for ligation, but none of these sites are restored after combining. Resulted pRTRA 7/3 vectors with *FLAG* gene are named pRTRA7/3 Flag2253, pRTRA7/3 Flag2251, pRTRA7/3 Flag1993 and pRTRA7/3 Flag1991 (Figure 22B). The inserted fragments were automatically in-frame with the c-myc and KDEL sequences. Clones bearing plasmids with the correct insertions were selected by restriction analysis and those clones were further used in a transient expression assay.

The transient expression assay in plant protoplasts is a powerful tool in plant molecular biology that allows quick screening and analysis of engineered proteins prior to stable transformation. Transient expression techniques enable the measurement of gene expression, as well as the analysis of protein stability and activity very shortly after DNA uptake. Therefore, a large number of samples can be analysed in a short period of time. The delivery of foreign genes into protoplasts is carried out by a PEG mediated transformation protocol (Reidt et al. 2000). The four different FLAG expression cassettes (Figure 18) were checked by transient expression in tobacco protoplasts. Protoplast were harvested, solubilised in SDS-sample buffer and analysed by PAGE and Western blotting analysis (see Materials and Methods). Proteins were detected by anti c-myc antibody (Figure 24A, B) and by anti FNT rabbit serum (Figure 23C, D). With anti FNT rabbit serum all four proteins were detected independent of the presence of c-myc tag. Samples were analysed at reducing and non-reducing conditions to check polymerisation of protein monomers *via* disulphide bridge formation of the cysteins contained in the C-terminal part of FLAG. The blots after separation at reducing (Figure 23A, C) and non-reducing conditions (Figure 23B, D) look identical. The proteins were of monomer size under both conditions. Polymerisation of proteins under non-reducing conditions was not detectable.

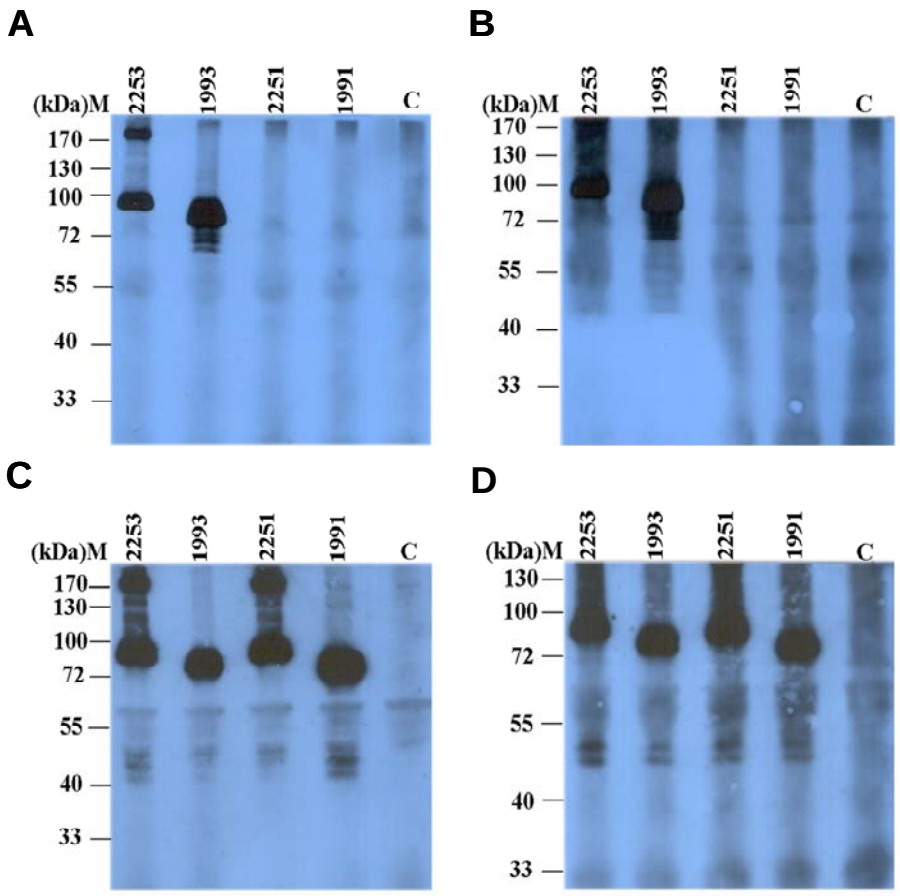


Figure 23 Western blot analysis of transiently expressed *FLAG* gene variants in tobacco protoplasts. (25 µg TSP/lane) Total protein extracts from protoplasts transformed with pRTRA7/3 Flag2253, pRTRA7/3 Flag1993, pRTRA7/3 Flag2251, pRTRA7/3 Flag1991 were separated under reducing (A, C) and non-reducing (B, D) conditions. Proteins were detected using monoclonal mouse c-myc tag antibodies, sheep anti mouse POD conjugate and ECL (A, B) and using rabbit anti N-terminus FLAG serum, donkey anti rabbit POD conjugate and ECL (C, D).

2.3.1.3 Cloning of *FLAG* gene expression cassettes into the binary vector pCB 301 Kan and transient expression by agroinfiltration

The production of transgenic plants is performed *via Agrobacterium*-mediated transformation. For this purpose, the expression cassettes have to be cloned into a shuttle vector providing border sequences and a resistance marker for the selection of transformed plants. The resulting shuttle vector then is introduced into a suitable *Agrobacterium tumefaciens* strain. The *FLAG* gene constructs in the four different expression cassettes were excised using HindIII restriction sites and cloned into the binary vector pCB301 Kan (Figure 24). The vector pCB301-Kan (Scheller et al. 2001) is based on the vector pCB301 (Xiang et al. 1999). This plasmid contains sequences encoding the npt III gene conferring kanamycin resistance in bacteria, the RK2 replication origin and a T-DNA region bordered by LB and RB and bearing the

kanamycin resistance (neomycin phosphotransferase II gene) as a plant screenable marker.

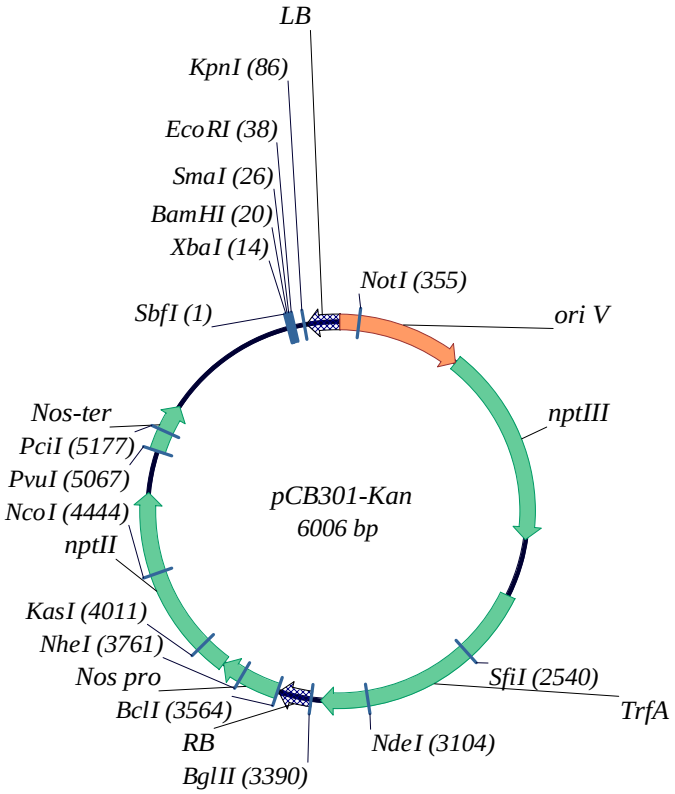


Figure 24. Scheme of the pCB 301 Kan binary vector.

Plasmids with target genes were transformed into *Agrobacterium tumefaciens* pGV 2260 cells (Deblaere et al., 1985) by electroporation and selected on kanamycin containing medium. To verify functionality of the cloned genes the transient expression by agroinfiltration (Kapila et al. 1997) was used because of its speed, convenience, and flexibility. Most of the introduced plasmid DNA remains extra-chromosomal during the transient assay (Werr et al. 1986) and, consequently, gene activity is not biased by position effects, as observed in stably transformed plants.
Clones which contain insertions of target genes were chosen for subsequent transient expression by agroinfiltration into tobacco leaves (see Materials and Methods).
PAGE and Western blotting experiments were performed to analyse expression levels of recombinant proteins in transiently expressed tobacco leaves (Figure 25). Proteins of expected size could be detected by the N-terminus specific antibodies. No reaction could be seen with the negative controls. The expression levels are different, but this could be also explained by the agroinfiltration method itself. The tests performed in tobacco leaves show successful expression of all four constructs in binary vectors as an important prerequisite for the production of stably transformed transgenic plants. The clones highlighted by arrows are chosen for stable transformation.

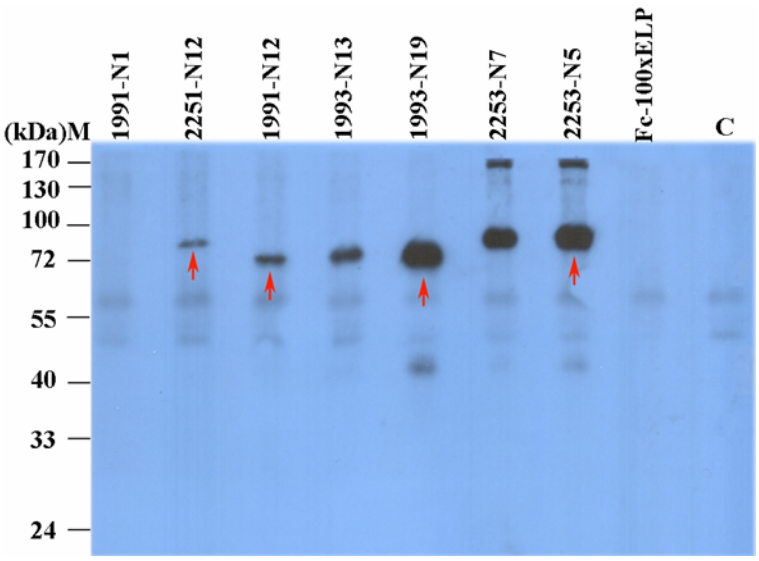


Figure 25. Western blot analysis of FLAG proteins, expressed after agroinfiltration in tobacco leaves. (25 µg TSP/lane).
FLAG proteins from different constructs (pCB301 Flag2253, pCB301 Flag1993, pCB301 Flag2251, pCB301 Flag1991) and different pCB-clones were separated at reducing conditions and detected using rabbit anti N-terminus FLAG serum, donkey anti rabbit POD conjugate and ECL. Negative controls: C (non-infiltrated leaf), Fc-100xELP.

2.3.2 Expression of spider silk and spider silk-ELP fusion proteins in transgenic plants

Potato tubers have a great potential in biofarming for production of foreign proteins. Specialised as a storage organ they can be used to accumulate large amounts of protein. Agricultural practice to grow, harvest, store and process potato tubers for starch production is well established and can be adapted for protein production. As potato is grown in different type of climates all over the world, its potential for commercial exploitation for biofarming is obvious. In addition, potato is the first of these agriculturally most important plant species which has been tested for applicability as a biofactory (Artsaenko et al. 1998). The cost of proteins produced in plants could be further decreased by developing a by-production system. Starch potatoes are the raw material for the production of technical grade starch. During this process soluble proteins are separated by heat treatment and sold as animal fodder. A by-production system of a recombinant protein in transgenic potato tubers would be useful for a bulk protein product with low monetary value, such as structural fibre proteins (Scheller et al. 2004).
J. Scheller and co-workers have reported the design of synthetic genes (SO1) to encode analogs of the MaSp1 protein which is composite of dragline silk of the golden orb weaver *Nephila clavipes,* and its expression in tobacco plants (Scheller et al. 2001). Later on, the plant-based production and purification of an SO1-ELP fusion protein has been shown (Scheller et al. 2004). In this construct SO1 is the synthetic 1.8 kb homologue of the *Nephila clavipes* MaSp1 cDNA, lacking the non-repetitive part at the 3` terminus. MaSp1-100xELP and MaSp2-100xELP were also designed by J. Scheller (unpublished data). The sequences of both proteins are obtained from *Nephila clavipes*. MaSp1 and

MaSp2 are two composites of dragline silk, which released from major ampullate gland of spider. Fusion of spider silk protein with the elastic biopolymer 100xELP (100 repeat of pentapeptide Val-Pro-Gly-Xaa-Gly) will provide easy separation of recombinant target protein from plant crude extract and, as shown for SO1-ELP in tobacco, increase the accumulation of the transgenic proteins. The three fusions of spider silk genes with the synthetic elastin gene 100xELP were introduced in an expression vector allowing ubiquitous expression in transgenic plants and ER retention (Figure 18). The *FLAG* constructs described above, never stably expressed so far, will be transformed into tobacco plants as most suitable plants to study basics of molecular farming.

2.3.2.1 Production and characterisation of transgenic potato and tobacco plants expressing the different spider silk-like proteins

After the feasibility of spider silk protein production in plant cells had been demonstrated by transient expression studies, the *Agrobacterium*-mediated DNA delivery system was used to create stable tobacco and potato transformants.
The pCB301 Kan based vectors, containing genes of different spider silk-like proteins (Figure 18) were transferred into *Agrobacterium tumifaciens* strains pGV 2260S and pGV 2260 (Deblare et al. 1985) by electroporation. Starch potatos (*Solanum tuberosum* cv. Albatros) and tobacco (*Nicotiana tabacum* cv. SNN) leaf discs were transformed according to (Zambriski et al. 1983). After *Agrobacterium tumefaciens* mediated leaf disc transformation, transgenic plants were selected on medium containing 100 mg/L kanamycin. Regenerated plants containing the transgene were selected for further investigation after Western blot analysis with anti c-myc monoclonal antibody (Munroe & Pelham 1987). Transgenic MaSp1-100xELP, MaSp2-100xELP and SO1-100xELP potato plants and transgenic pCB301 Flag2253, pCB301Flag1993, pCB301Flag2251 and pCB301Flag1991 tobacco plants were regenerated from leaf-discs transformed with the *Agrobacterium*-clones containing the genes of spider silk and spider silk-like proteins under control of the constitutive CaMV 35S promoter.

2.3.2.1.1 Production of transgenic tobacco plants expressing different FLAG variants.

Agrobacterium strains with the 4 different constructs (pFlag2253, pFlag1993, pFlag2251, pFlag1991, see chapter 2.3.1.3.) in pCB301 have been used for the stable transformation of tobacco plants. For every construct a number of high expressor plants could be selected (Table 9).

Table 9. Expression of FLAG protein in transgenic plants: constructs, number of analysed resistant plants and number of plants expressing the transgene

Constructs	**Number of analysed Kan - resistant plants**	**Number of producer plants**	**Number of plants with high expression level**
Flag_2253 (Tobacco sp. Samsun NN)	50	18	18
Flag_1993 (Tobacco sp. Samsun NN)	50	30	30
Flag_2251 (Tobacco sp Samsun NN)	40	27	18
Flag_1991 (Tobacco sp Samsun NN)	50	31	14

As shown in Figures 26, 27, 28 and 29 proteins of expected size were expressed in several lines.

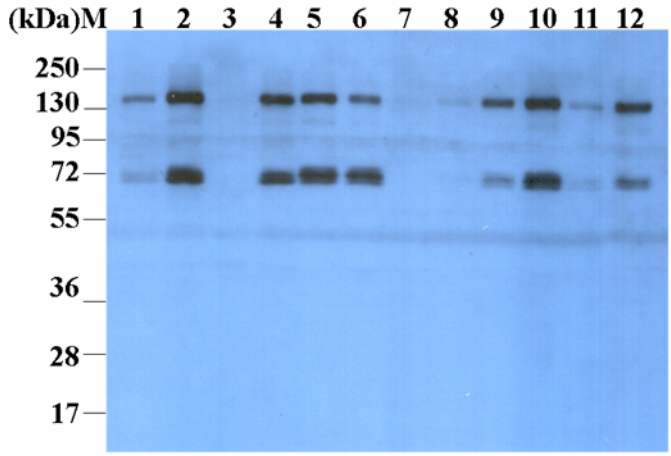


Figure 26. Expression of Flag2251 (N-terminus-repetitive ensemble repeat-C-terminus) in transgenic tobacco plants (TG0) shown by Western blot. Leaf extracts (20 µg TSP/lane) from TG0 lines were separated at reducing conditions and detected using rabbit anti N-terminus FLAG serum, donkey anti rabbit POD conjugate and ECL.

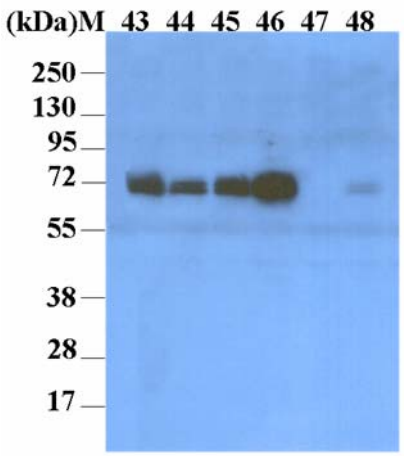


Figure 27. Expression of Flag1991 (N-terminus-repetitive ensemble repeat) in transgenic tobacco plants (TG0) shown by Western blot. Leaf extracts (20 µg TSP/lane) from TG0 lines were separated at reducing conditions and detected using rabbit anti N-terminus FLAG serum, donkey anti rabbit POD conjugate and ECL.

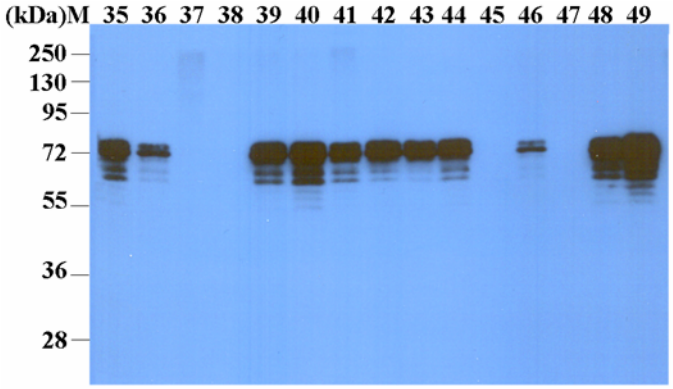


Figure 28. Expression of Flag1993 (N-terminus-repetitive ensemble repeat)-c-myc in transgenic tobacco plants (TG0) shown by Western blot. Leaf extracts (20 µg TSP/lane) from TG0 lines were separated at reducing conditions and detected using monoclonal mouse anti c-myc tag antibodies, sheep anti mouse POD conjugate and ECL.

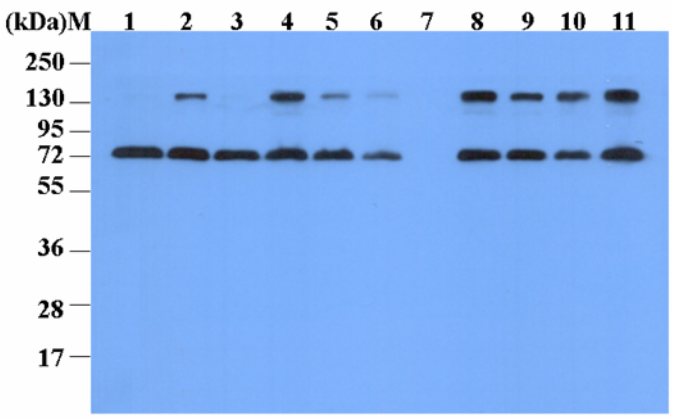


Figure 29. Expression of Flag2253 (N-terminus-repetitive ensemble repeat-C-terminus)-c-myc in transgenic tobacco plants (TG0) shown by Western blot. Leaf extracts (20 µg TSP/lane) from TG0 lines were separated at reducing conditions and detected using monoclonal mouse anti c-myc tag antibodies, sheep anti mouse POD conjugate and ECL.

In all plants expressing FLAG proteins with C-terminus a second band of higher molecular weight was visible. Further experiments (non-reducing conditions, stronger reduction) showed that this is not due to multimerisation *via* disulphide bridges (data not shown). Best producer plants were selected and will be further propagated. The plants designed and partially characterised here will be the base of an extended purification and characterisation programme.

2.3.2.1.2 Characterisation of transgenic potato plants expressing different spider silk-like proteins.

Protein production in leaves of the transgenic potato plants was examined by Western analysis (Figure 30, Figure 31, Figure 32).

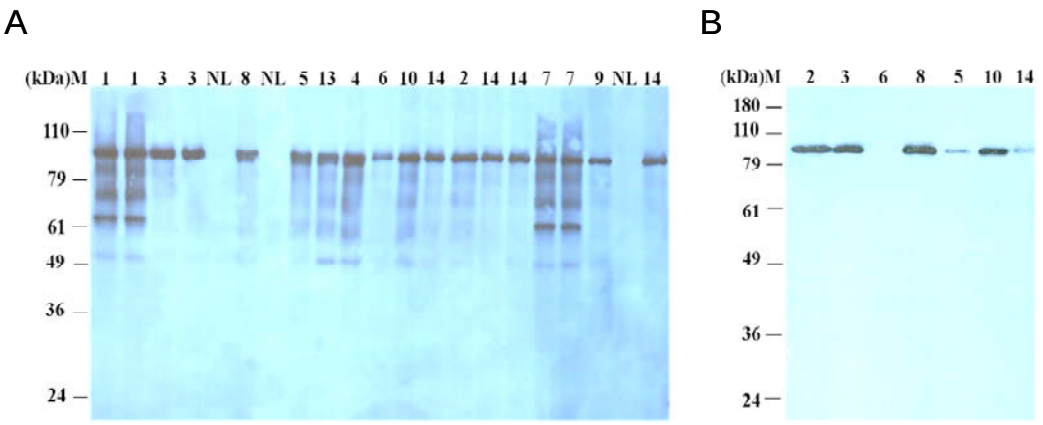


Figure 30. Expression of MaSp1x100ELP in transgenic *Solanum tuberosum* cv. Albatros shown by Western blot. (A) MaSp1x100ELP leaf extracts (5 µg TSP/lane) and (B) MaSp1-ELP tuber extracts (5 µg TSP/lane) from TG0 lines were separated at reducing conditions and detected by monoclonal mouse anti c-myc tag antibodies, sheep anti mouse POD conjugate and ECL. Numbers represent independent TG0 lines; NL: not loaded.

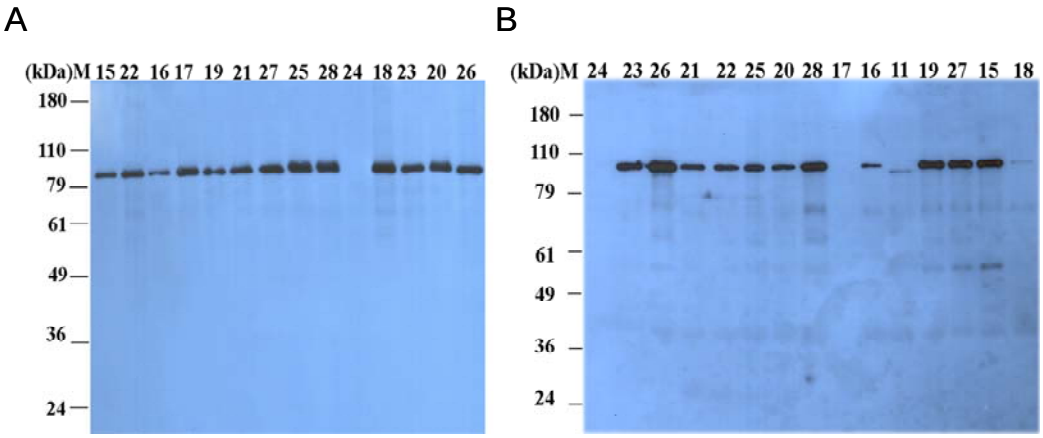


Figure 31. Expression of MaSp2x100ELP in transgenic *Solanum tuberosum* cv. Albatros shown by Western blot. (A) MaSp2x100ELP leaf extracts (5 µg TSP/lane) and (B) MaSp2-100xELP tuber extracts (5 µg TSP/lane) from TG0 lines were separated at reducing conditions and detected by monoclonal mouse anti c-myc tag antibodies, sheep anti mouse POD conjugate and ECL. Numbers represent independent TG0 lines.

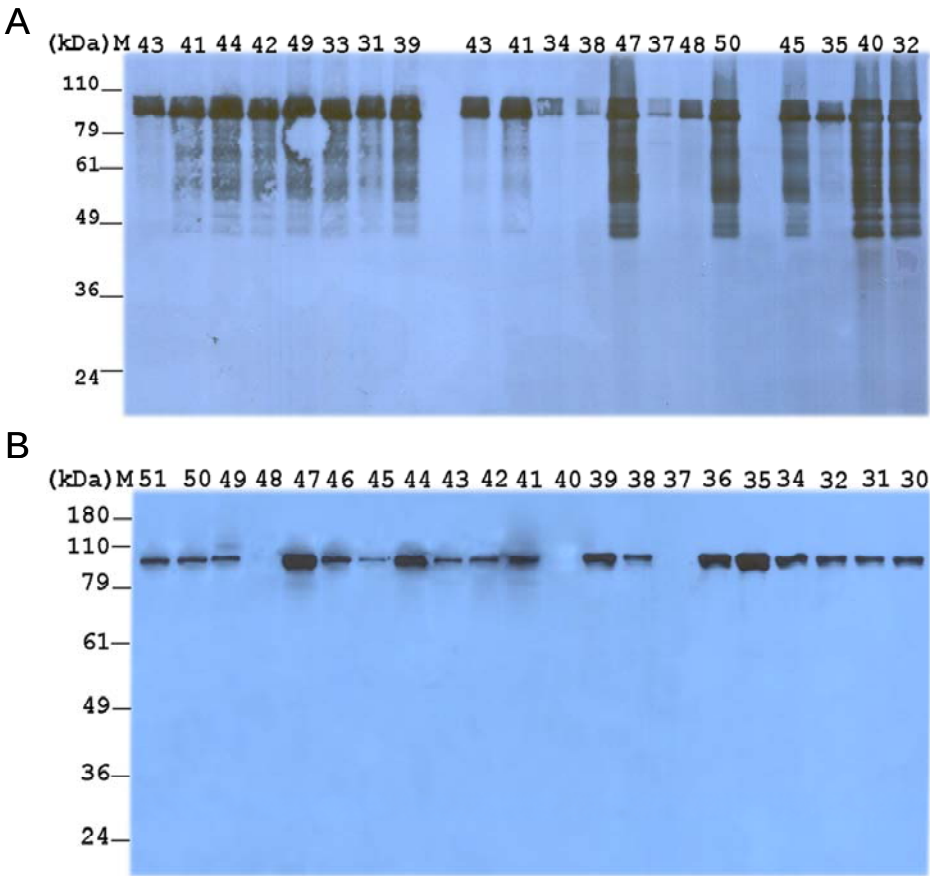


Figure 32. Expression of SO1x100ELP in transgenic *Solanum tuberosum* cv. Albatros shown by Western blot. (A) SO1x100ELP leaf extracts (5 µg TSP/lane) and (B) SO1-ELP tuber extracts (5 µg TSP/lane) from TG0 lines were separated at reducing conditions and detected by monoclonal mouse anti c-myc tag antibodies, sheep anti mouse POD conjugate and ECL. Numbers represent independent TG0 lines; NL: not loaded.

The number of regenerated Km-resistant plants and the number of transgene expressing plants are shown in Table 10. The TG0 lines showing higher recombinant protein expression were selected and propagated for greenhouse experiments. Exact quantification of the expression level was not done with TG0 plants, but known amounts of TSP are applied to the lanes. For all constructs stable expression of the spider silk proteins in starch potato was shown in several lines. The sizes of the recombinant plant-produced proteins were as expected. In many cases expression levels in leaves and tubers are comparable (Figure 31A, B, lines 16, 20, 21, 22, 27, 28), in several cases expression in leaves is higher than in tubers (Figure 31A, B, lines 17, 18, 25) and in other cases expression in tubers is higher than in leaves (Figure 31A, B, lines 19, 23, 26).
As shown here, the relation between tuber and leaf expression is different in the individual lines. The next question to be addressed here is the expression level in single tuber of one line. For this purpose the line SO1-100xELP, line 41 was propagated *in vitro* and then 18 plantlets were grown in pots in the greenhouse and one tuber from each pot was analysed by Western blot (Figure 33).

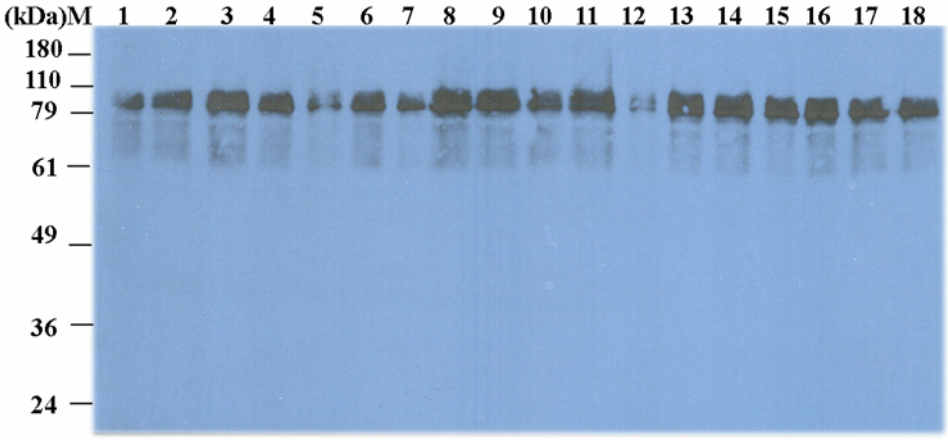


Figure 33. Expression of SO1x100ELP in tubers of 18 different clones of the line SO1x100ELP 41 shown by Western blot. Every number represents one tuber/ clone grown in individual pots. Extracts were separated at reducing conditions and detected by monoclonal mouse anti c-myc tag antibodies, sheep anti mouse POD conjugate and ECL. 25 µg of protein loaded onto the gel.

Medium size tubers were used in these experiments. As shown in Figure 33, the expression level is quiet different from pot to pot. Further preliminary analyses adressed the questions, if the transgene expression influences tuber morphology and yield. In greenhouse experiments the diameter of tubers, the morphology and the tuber yield was measured in different lines of the constructs MaSp1-100xELP, MaSp2-100xELP and SO1-100xELP. The results are presented in Figures 34A, B, C and Figure 35. In the transgenic lines of the constructs MaSp1-100xELP and MaSp2-100xELP the diameter of tubers differs from line to line. High variability in the lines is represented by high standard deviation values (Figure 34A, B, C). This high variability in size is also documented in Figure 35. In transgenic lines of all three constructs (MaSp1-100xELP, MaSp2-100xELP and SO1-100xELP) small tubers occur, whereas there are also tubers of normal size. Widtype tubers do not differ as much in size. The yield of tubers from greenhouse plant strongly depends on growth conditions and seasonal differences. This explains the differences in wildtype yields between the different cultivations (Figure 36). There are always transgenic lines showing the same tuber yield as the wildtype (MaSp1-100ELP 14, MaSp2-100xELP 26, SO1-100xELP 33 and SO1-100xELP 44), but several lines have a much lower yield (MaSp1-100ELP 8, SO1-100xELP 39).

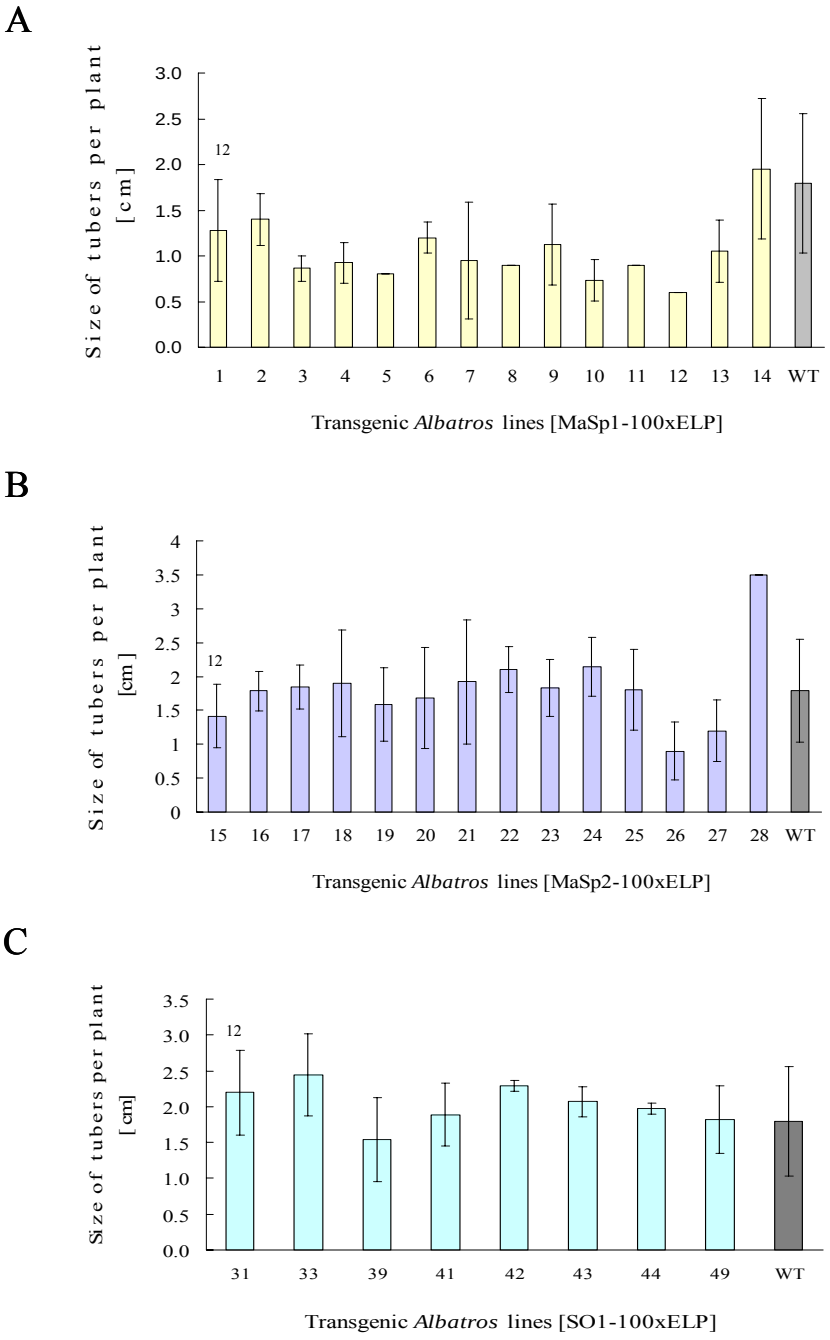


Figure 34. Variation of tuber sizes. (A) within 14 different TG0 lines of MaSp1-100xELP construct. (B) within 14 different TG0 lines of MaSp2-100xELP construct. (C) within 14 different TG0 lines of SO1-100xELP construct. Plants were grown at greenhouse conditions. The number above each bar represents harvested potato tuber numbers. Bars marked by stars represent tubers with equal size.

Figure 35. Tuber size and morphology in individual transgenic lines of MaSp1-100xELP, MaSp2-100xELP and SO1-100xELP compared to wildtype. Plants are grown in pots in the greenhouse at the same time.

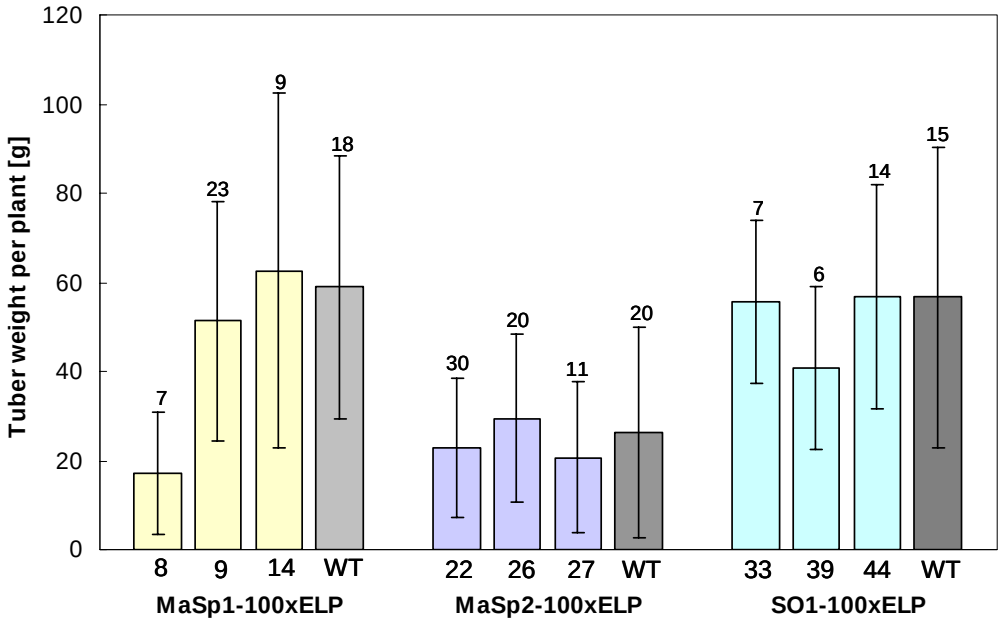


Figure 36. Tuber weight per plant in individual transgenic starch potato lines expressing spider silk-ELP fusion proteins. Plants were grown in the greenhouse. The different wildtype controls were grown at the same time as the related transgenic lines, respectively.

The growth of potatos at greenhouse conditions leads to rather limited knowledge, how the transgene expression influences growth, yield and other parameters. Therefore, further analyses have then been done with this material in a field trial performed in 2005. The transgenic potato lines described above were used in this field trial as a major prerequisite. The first general goal in this trial was to a produce recombinant spider silk derivative in starch potato plants and to purify the transgenic protein. To this aim selected transgenic lines were amplified and healthy tubers were produced in the greenhouse by a potato breeding company (NORIKA Groß Lüsewitz). The main tasks of this trial were production of MaSp1-100xELP- und MaSp2-100xELP potatoes in the field and extraction and purification of these proteins for the first time. One main presumption for the use of this expression and production system is the stability of the spider silk proteins during storage of the harvested potatoes for at least a few months. Unexpectedly, in starch potatoes the spider silk proteins were not stable during storage. Even after a few weeks at 4°C the transgenic spider silk derivatives were cleaved by proteases (C. Münnich and U. Conrad, unpublished). Here, a basic presumption of the starch co-expression system was not fulfilled after these field studies. In addition to the losses during storage these proteolytic activities mainly prevented the development of a suitable and cheap purification system. Alternatively, tobacco leaves and seeds are now in the focus of research to establish a practically applicable plant based production system.

Table 10. Expression of spider silk proteins and spider silk-ELP fusion proteins in transgenic plants: constructs, number of analysed resistant plants and number of plants expressing the transgene.

Construct	Number of analysed Kan - resistant plants	Number of producer plants*	Number of plants with high expression level
Spidroin1-ELP (MaSp1x100ELP, *Solanum tuberosum* cv. Albatros)	62	13	13
Spidroin2-ELP(MaSp2x100ELP, *Solanum* tuberosum cv. Albatros)	58	15	15
Synthetic spidroin1-ELP (SO1x100ELP, *Solanum tuberosum* cv. Albatros)	67	15	15

*Producer plants: plants accumulating recombinant spider silk protein in higher amount.

2.4 Production of Fc-100xELP fusion proteins in plants

The choice of a suitable system for recombinant protein production relies on the abilities of the heterologous host to post-translationally modify the protein of interest so that it is identical or very similar to the native protein. In this regard, plant cells are generally able to properly fold, post-translationally modify, and otherwise process recombinant proteins in a manner that is similar to that of animal systems. It is known that all spider silks are polymers of high molecular weight with sizes ranging from 215-320 kDa (Mello et al. 1994; Candelas et al. 1981). The recently obtained data on spider silk gene sequences demonstrated that mRNA size of studied spider silk genes is rather big (up to 15 kb, Hayashi et al. 1998). Most recombinant studies on spider silk proteins are performed with partial cDNA sequences due to the absence of full-length clones of spider silk proteins.
The maximal size of spider silk-like polypeptides to be expressed in plants is limited by the stability of these highly repetitive sequences in *Agrobacterium tumefaciens*. Therefore it is necessary to enhance the size of recombinantly produced proteins post-translationally. Furthermore, spider silks as well as silks produced by *Lepidopteran* known to contain from two to three cysteine residues at their C-terminus or N-terminus. These cysteine residues were shown to be involved in multimerisation of protein molecules *via* disulphide bond formation (Sponner et al. 2004). The other critical point in recombinant protein production in heterologous systems is to develop easy purification methods. To achieve these two goals as post-translationally dimerisation of structural

repetitive protein molecules and easy purification we used in our study the fusion of 100xELP to the Fc domain of human immunoglobulin G (Figure 37). Due to a signal peptide the fusion protein is translated at the ER associated ribosomes and protruded into the endoplasmic reticulum, where the dimerisation *via* disulphide bridges takes place. Furthermore the disulphide bond between cysteine side-chain sulfurs has been described as a force for stabilising the protein structure for a prolonged period.

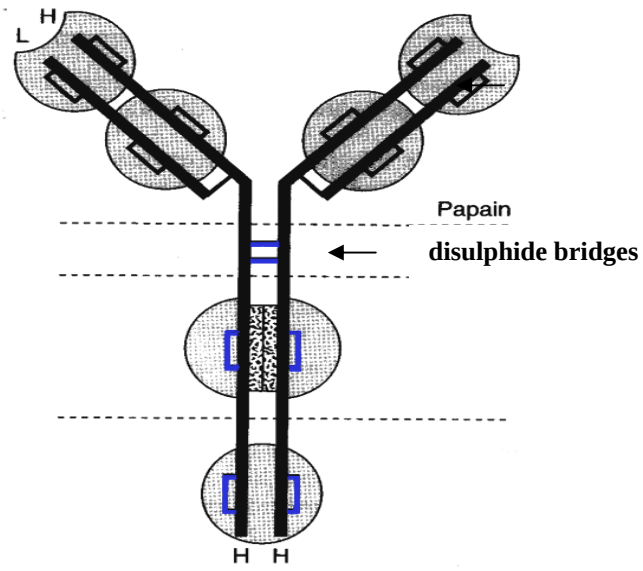


Figure 37. Schematic picture of an immunoglobulin G molecule.

2.4.1 Construction of plant expression cassettes containing Fc-100xELP encoding sequences

pRTRA 7/3-sgp130-Fc-100xELP (kindly provided by Dr. J. Scheller) was used as a template for cloning of the plant expression construct. The digestion of plasmid pRTRA 7/3-sgp130-Fc-100xELP with restriction enzymes BamHI and EcoRI and subsequent ligation of this vector with T4 DNA ligase resulted in plasmid pRTRA-Fc-100xELP (Figure 38).
The plasmid pRTRA 7/3-Fc-100xELP was digested with HindIII and the fragment containing the cassette CaMV 35S promoter/legumin signal/c-myc tag/Fc/100xELP/KDEL/CaMV 35S Terminator was cloned into pCB301 Kan (Figure 24). Resulting plasmid was transformed into *Agrobacterium tumefaciens* pGV 2260 and pGV 2260S (Deblaere et al. 1985) for subsequent transformation into tobacco and potato plants, respectively.

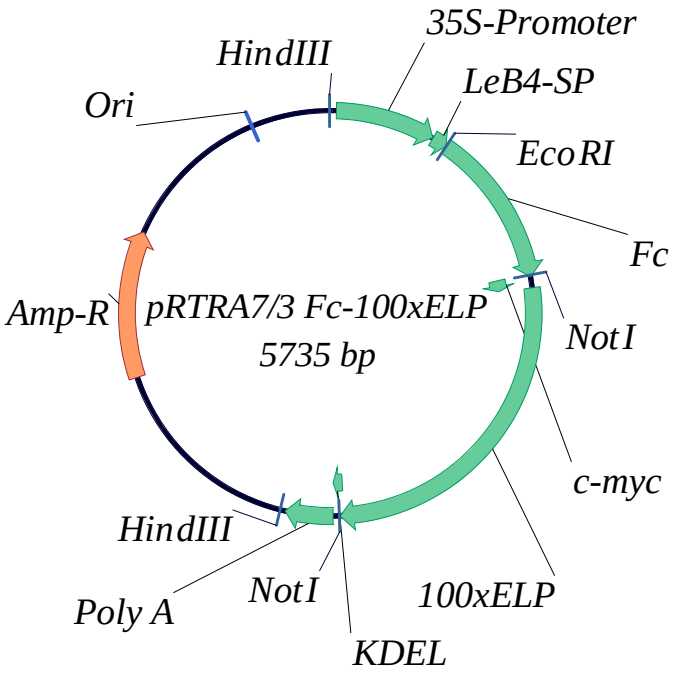


Figure 38. pRTRA 7/3 vector containing Fc-100xELP encoding sequences fused at the N-terminal part with LeB4 signal peptide and at the C-terminal part with KDEL. The peptide containing c-myc epitope is located between Fc and 100xELP.

2.4.2 Production and characterisation of Fc-100xELP expressing tobacco and potato plants

Both in tobacco and starch potato background Fc-100xELP expressing plants were selected. The molecular weight of the expressed transgenic protein was as expected (Figures 39, 40). Generally, the expression level in the different transgenic lines was comparable in tobacco and potato leaves. Seeds from all tobacco expression lines were stored and then further propagated for purification. The lines 5, 7 and 9 from the transgenic potato collection have been chosen and kept as *in vitro* lines.

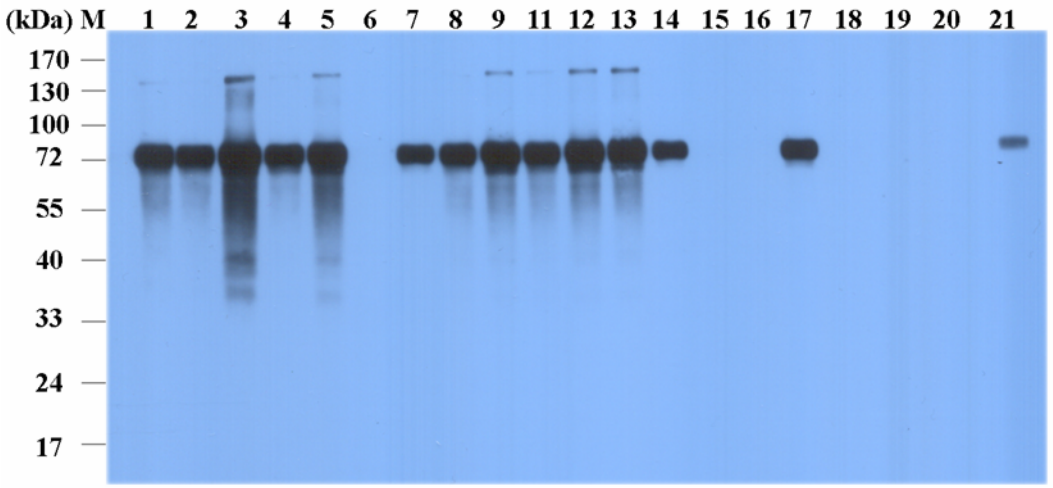


Figure 39. Expression of Fc-100xELP in transgenic tobacco plants (TG0) shown by Western blot. Leaf extracts (5 µg TSP/lane) from TG0 lines were separated at reducing conditions and detected using monoclonal mouse anti c-myc tag antibodies, sheep anti mouse POD conjugate and ECL.

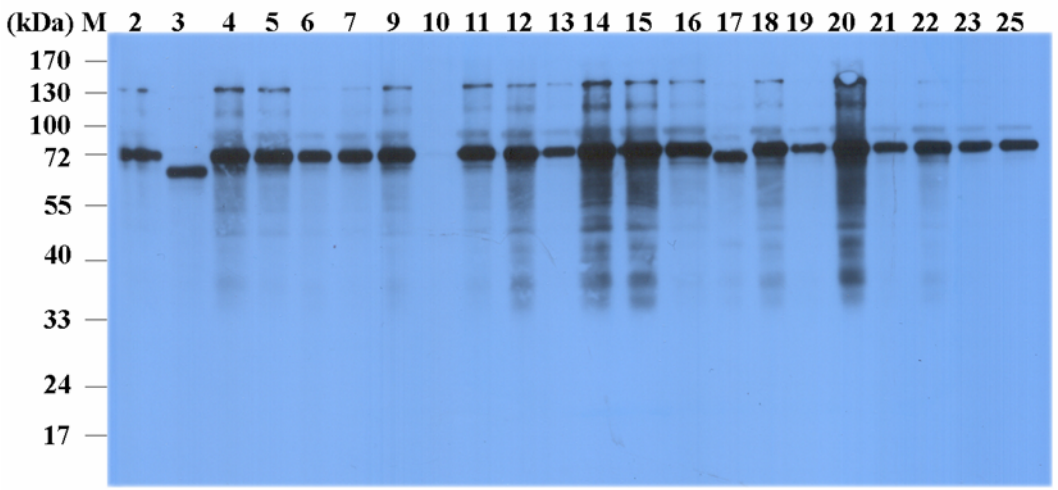


Figure 40. Expression of Fc-100xELP in transgenic starch potato plants (cv. Albatros, TG0) shown by Western blot. Leaf extracts (5 µg TSP/lane) from TG0 lines were separated at reducing conditions and detected using monoclonal mouse anti c-myc tag antibodies, sheep anti mouse POD conjugate and ECL.

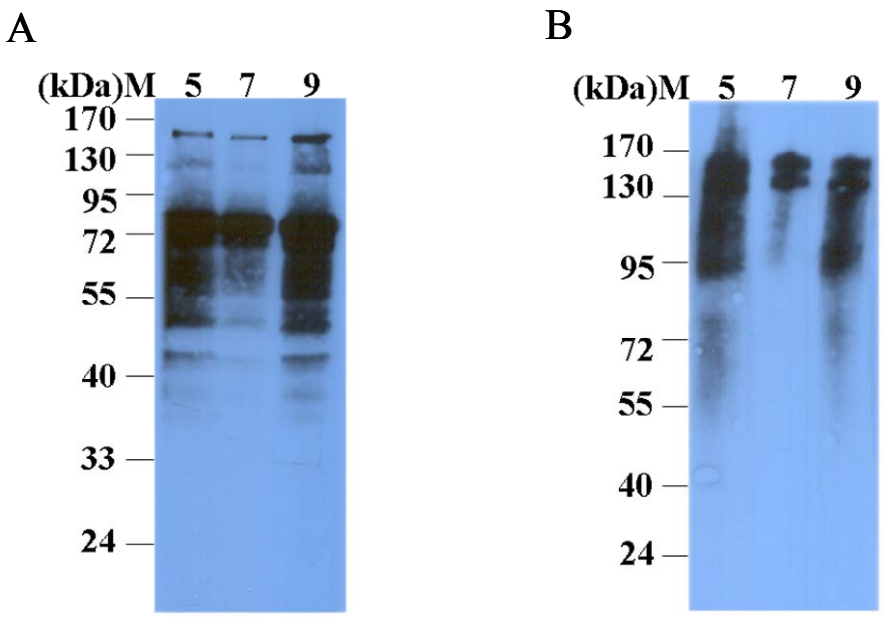


Figure 41. Western blot analysis of Fc- 100xELP from potato lines 5, 7 and 9, separated (A) at reducing conditions and (B) at non-reducing conditions and detected using monoclonal mouse anti c-myc tag antibodies, sheep anti mouse POD conjugate and ECL. (5 µg TSP/lane).

The separation of Fc-100xELP from potato leaves at reducing and non-reducing conditions shows, that indeed dimers are formed and that could be cleaved by β-mercaptoethanol. The double bands visible in Figure 41 occur probably because of glycosylations. Glycosylation should be expected for the Fc parts.
Fc-100xELP was purified from re-seeded tobacco lines *via* protein A affinity chromatography. The protein was highly soluble in water, but not soluble in HFIP (hexafluorisopropanol). The protein is under study now according mechanical properties (F. Junghans, U. Spohn, Fraunhoferinstitut für Werkstoffmechanik Halle, unpublished).

3 Discussion

Methods of genetic engineering of recombinant proteins and plant transformation techniques have been used to create tobacco plants expressing the spider silk FLAG protein.

A strategy developed for the production of transgenic plants expressing the FLAG protein included several stages: (i) cloning of the partial *FLAG* cDNAs, (ii) demonstration of the feasibility of the FLAG protein production in plant cells by transient expression system and (iii) construction of plant expression vectors for ubiquitous expression of FLAG protein in the ER of plant cells and *Agrobacterium*-mediated transformation of tobacco plants. It is expected that strategy of *FLAG* gene design with its N- and C- termini can be applied for other known silk genes, thus opening new opportunities for investigation of the structure and properties of recombinant spider silk proteins.

A prerequisite for this experiment (expression of FLAG with its own N-terminus and C-terminus) was the availability of cDNA clones encoding repetitive parts, non-repetitive N-terminus and C-terminus of *FLAG*-gene (kindly provided by C.Y. Hayashi). The recombinant gene encoding FLAG's N-terminus-repetitive unit- C-terminus was assembled into expression cassette under plant expression regulator elements (see Figure 18). In order to target recombinant proteins into the ER lumen KDEL peptide was fused at the C-terminus of FLAG, which was expected to act simultaneously with putative signal peptide (SP) located at the 5'-end of N- terminus of FLAG (see Figure 1B).

In first experiments retention of our recombinant proteins into the ER we expected to detect protein polymerisation events by comparative analyses of target protein in reducing and non-reducing conditions. These expectations are based on the presence of two cysteine residues at the C-terminus peptide of FLAG. C-terminus of dragline silks were previously shown to be involved into cross-linking of spider dragline silk protein monomers into polymers (Mello et al. 1994; Sponner et al. 2004). As shown in Figure 23, dimmers or polymers depending on disulphide bridges could not be detected after transient expression. Stable transformants under study.

In order to study possible regulatory roles of the C-terminus of FLAG in self-assembling processes two *FLAG*-gene variants were designed lacking C-terminal peptide.

Thus a comparison of effects on protein polymerisation, through expression of the *FLAG*-gene variants lacking C-terminal peptide with cysteine residues is of interest especially regarding dissection of the possible function of C-terminus peptide of FLAG in cross-linking of protein monomers.

Inserting a tag peptide into a protein facilitates the characterisation of that protein when antibodies to the protein are not available. Recent techniques incorporate a peptide tag into a recombinant protein to aid protein purification or isolation of that protein. Antibodies recognising the peptide tag facilitate purification and/or isolation. It is important, that the addition of the tag to the C-terminus of a protein does not negatively affect the conformation of that protein. One method to predict whether a particular insertion of tag will alter secondary structure of the target protein is to use computer modelling to reconstruct the secondary structure of the gene product of interest. Another method to identify tag insertion influence is to insert nucleic acid encoding tag into a nucleic acid encoding the protein and test the protein for activity or for pI and total net charge of protein. In our case the insertion of c-myc tag into the FLAG protein sequence

at its C-terminus have been calculated to alter pI of protein from 5.05 to 4.73 (see Table 8 in Chapter 2). Taking into consideration that total net charge of protein molecules is critical for self-assembling processes occurring in the environment of spinning duct *in vivo*, we decided to avoid c-myc tag in two variants of *FLAG* gene constructs.

FLAG is a protein polymer which is the main fibrous protein of flagelliform silk. Flagelliform silk build up capture spirals in spider web and is characterised to be elastic and tough. Comparison of primary structures of FLAG to other known elastic proteins from different sources: elastin from human, resilin from insects, gluten from wheat revealed similar structure as G-P- conformers which are shown to form β-turn/β-spirals (Urry et al. 1988). So the study on FLAG protein could give us insights into the basis of elasticity in silk and silk-like proteins.
FLAG gene architecture has been characterised. Among several partial cDNAs known for FLAG two are encoding FLAG's N-terminus and C-terminus, whereas others encode repetitive parts of FLAG. Previous studies on *Bombyx mori* silk genes reported the importance of non-repetitive N- and C-termini of silk protein in self-assembling processes *in vivo* (Jin and Kaplan 2003) in formation of outer edges of micelles. The micelle is a phase likely to be present at relatively low silk protein concentrations. The roles of the non-repetitive N- and C-terminal domains, and whether the N-terminus is present in the silk fiber or removed during the fiber assembly (as are terminal domains of procollagen), for example, are unknown.
Importantly, there are a number of studies on expression of spider silk repetitive regions (Prince et al. 1995; Fahnestock et al. 1997; Scheller et al. 2001) and repetitive region with non-repetitive C-terminus (Huemmerich et al. 2004; Arcidiacono et al. 1998), but so far no reports on expression of spider silk proteins with its N-terminus. This can be explained due to the more growing interest of scientists to dragline silk proteins because of its outstanding mechanical properties. Also the main limitation was the lack of N-terminus sequences due to the big size of genes, and its repetitiveness. Flagelliform silk protein was the only one where N-terminus sequence was published (Hayashi et al. 1998).
The architecture of *FLAG* gene is shown to be an exon/intron structure, where 13 exons are interspersed by introns with high homogeneity in size and sequence. Each exon is 440 amino acid in length and represented by one ensemble repeat, where a pentapeptide motif (GPGGX)n is followed by (GGX)n motif and spacer ending with GGX motif. Differences between exons appeared in slight variations in numbers of (GPGGX)n and (GGX)n motifs in range of 43 to 65 times and from 6 to 12 times, respectively. The spacer is known to be 28 amino acid in length, non-repetitive and highly conserved unit. Therefore, primary coding regions of the FLAG protein are homogeneous and selective splicing of mRNA transcripts would result in different sized protein domains that, in turn, could affect the protein's structure and function (Craig et al. 2002).
Here, we combined in one continuous gene expression cassette partial cDNA sequences of *FLAG* from two distal exons, encoding N-terminus of FLAG and its repetitive part and C-terminus of FLAG. Our constructed gene sequence encodes FLAG protein containing N-terminus (101aa), C-terminus (87aa) and repetitive part (530aa), where pentapeptide (GPGGX) motif appeared up to 51 times, followed by (GGX) motif which repeated up to 6 times and included one spacer motif ending with GGX motif and followed by

(GPGGX) motif again repeated up to 18 times. Schematically primary structure of our protein is represented in Figure 44 (see Figure 44, Appendix). Our native *FLAG* gene designs attempted to capture as much of the sequence regularity of the natural proteins as possible because we had, and still have, a limited understanding of structural importance of the sequence patterns. The construction of *FLAG* gene very close to the native one is a first step forward to mimic nature in producing biomaterials.

FLAG seems to be a suitable object to study correlation between the protein's size and structure. This argument can be broadened by recognising that the protein participates in at least three different phases.

In the lumen of the spider's gland the well studied dragline silk protein is stored in an extremely viscous phase that is resistant to shear forces resulting from the movements of the spider (Vollrath and Knight 1999). A second, liquid crystalline phase of lower viscosity apparently forms near the exit duct (Willcox et al. 1996; Vollrath and Knight 1999), and may pre-align molecules for the subsequent formation of the third, insoluble fibre phase. To produce this state, the mesogenic (liquid crystal forming) units, whether domains within the protein molecule, whole molecules or aggregates of one or more different protein molecules, must be amphiphilic, that is having a combination of hydrophobic and hydrophilic blocks or groups (Bini et al. 2004).

There is currently a gap in understanding how the distribution of hydrophilic and hydrophobic blocks and other features of the primary sequence of silk proteins determines secondary, tertiary and quaternary structure of the protein and in turn the liquid crystal phase diagram.

The spidroins (the principal protein components of spider silks) are constructed largely from highly repetitive AB block copolymers showing a regular alternation of hydrophobic and less hydrophobic blocks (Vollrath et al. 2005).

Protein domains are defined as autonomous units of folding and frequently, of function (Doolittle 1993).

In contrast to *Bombyx mori* silk heavy chain fibroin where clearly designated large hydrophobic domains are interspersed by small hydrophilic spacer domains (Bini et al. 2004), flagelliform silk protein of *Nephila clavipes* does not contain β-sheet forming crystalline regions as poly-Ala which is known to form hydrophobic regions in dragline silk proteins or Gly-Ala stretches found in fibroins (Hayashi et al. 1999). The repetitive region of flagelliform silk is a tandemly repeated array of glycin rich pentapeptide motif interspersed by hydrophilic spacers. The alternation of hydrophilic and hydrophobic blocks in silk proteins has been suggested to form lamellar liquid crystal phases (Knight 2002) by phase separation of the hydrophobic from the hydrophilic blocks. The sequence information in the glycine-rich sequences may be involved in assembly of the critical liquid-crystalline phase.

Silk proteins appear to have extremely large molecular weights, probably the largest for any class of proteins. This is not accidental as mechanical properties of polymers depend on molecular size as in other liquid crystalline systems.

It could be speculated as more bigly the size of spider silk proteins as easily could higher concentrated protein solution with liquid crystalline phase be obtained. Also possibly hydrophobic blocks can not be strictly characterised by presence of poly-Ala blocks, because flagelliform silk protein from *Nephila clavipes* does not contain any poly-Ala

blocks, but its proline containing glycine-rich domains were theoretically calculated to comprise hydrophobic residues in one ensemble repeat within range of 6 to 15 %.
Selective splicing of FLAG mRNA would result in different sized protein domains, but there can not be a big difference in primary amino acid sequences of the resulting protein molecules, since all exons are more or less homogeneous. Thus, the FLAG protein structure and function could be affected mainly by the size of protein molecules. This hypothesis could be easily tested by designing gene constructs encoding FLAG's repetitive part in tandem repeated array of repetitive units. The *FLAG* gene constructs with different length can be expressed in heterologous systems which can maintain biosynthesis and proper processing of high molecular weight proteins with high repetitive structure. First of all it is necessary to find an appropriate host for heterologous expression, where highly repetitive gene can be maintained through generations. Recent works on the production of new biopolymers in transgenic plants were reviewed by (Scheller and Conrad 2005). The results shown by Scheller and co-workers (Scheller et al. 2001) are a preliminary development towards high level production of spider silk proteins or their derivatives in the ER of plant cells.
Expression of proteins in engineered cells is now a mature technology, and its scope and limitations are well understood:

a) Small proteins (i.e. <30 kDa) are easier to express than large multidomain proteins.
b) Folding of large protein molecules can also be a challenge.
c) Product heterogeneity is frequently a problem, caused by uncontrolled processing of the nascent polypeptide in the cell.
d) The overexpression of proteins that are toxic to the cell can be difficult.

We believe that transgenic plants are powerful protein factories that can overcome the problems mentioned above. Therefore, we constructed stable transgenic tobacco lines to express various native *FLAG* gene variants ranging from 1991 to 2253 base pairs. As well as providing a suitable approach for structural studies of plant-derived spider silk proteins, recombinant FLAG derivatives are of biotechnological significance.
In order to post-translationally increase the size of spider silk proteins we developed a Fc-100xELP fusion protein, where Fc is a part of human immunoglobulin IgG and has several cysteine residues which would result in two intermolecular S-S bridges. The conformation of the fusion protein in aqueous-based solution seems to be globular (Junghans and Sponner, unpublished). In addition, genetic engineering has one technical limitation: it is only applicable the preparation of a fusion protein that is in tandem from the N-terminus to the C-terminus, such as the scFv (Artsaenko et al. 1995) or SO1-100xELP fusion protein (Scheller et al. 2004).
It follows that we need a method that allows the linking of protein domains not only in a linear form, as in fusion proteins, but also in more complicated assemblies.
Transglutaminase, which can catalyse acyl transfer reactions between Gln and Lys residues (Folk 1983), provides an opportunity to assemble protein molecules as opposed to conventional techniques. This approach is under study in the phytoantibody group (Dr. Conrad U., IPK).

4 Material and methods

4.1 Material

4.1.1 Plant material

Nicotiana tabacum, cv.Samsun NN
Solanum tuberosum cv. Solara and Albatross
Nicotiana plumbaginifolia used for protoplast preparation

4.1.2 Strains

4.1.2.1 *Agrobacterium tumefaciens*

C58C1Rf^r (pGV 2260 in C58C1) [Deblaere et al. 1985]
C58C1Rf^r (pGV 2261 in C58C1) [Deblaere et al. 1985]

4.1.2.2 *Escherichia coli*

1. TG1: *K12, Δ(lac-pro), supE, thi, hsdD5/FtraD36, pro* A^+B^+, *lac1*^{q}, *lacZΔM15.*
2. HB2151: *K12, ara, Δ(lac-pro), thi/F' pro* A^+B^+, *lac1*^{q}*ZΔM15*.(Hoogenboom et al. 1991)
3. XL1 Blue: *recA1, endA1 gyrA96 thi-1, hsdR17, supE44 relA1 lac[F´proAB lac1*^{q}*ZΔM15 Tn10 (Tet*^{r}*)]* (Stratagene)
4. BL21 (DE3): *pLysS (ampR) BF*$^-$*dcm, ompT, hsdS(rb*$^-$*mb*$^-$*), gal, y((DE3) (pLysS Cam1))*

4.1.2.3 Yeast

Pichia pastoris (KM 170)

4.1.3 Vectors

phagemid pIT1	amp^r (from MRC center, Cambridge, UK)
phagemid pIT2	amp^r (from MRC center, Cambridge, UK)
pRTRA 7/3	amp^r (Artsaenko et al. 1996)
pGEM-T	amp^r (Promega)

pET23a(+)	amp^r [(Studier and Moffatt et al. 1986); (Rosenberg et al. 1987) (Nowagen)
pCB301	kan^r (Scheller et al. 2001)
pPICZαB	Zeo^r (www.invitrogen.com)

4.1.4 Spider silk protein genes in pGEM-T vector

N-terminal *FLAG* gene sequence isolated from *Nephila clavipes* (C. Hayashi)
C-terminal *FLAG* gene sequence isolated from *Nephila clavipes* (C. Hayashi)
C-terminal *MaSp1* gene sequence isolated from *Argiope trifasciata* (C. Hayashi)
C-terminal *MaSp2* gene sequence isolated from *Araneus gemmoides* (C. Hayashi)

4.1.5 Primers (Metabion)

PCR1 primers

FOR Fl_N-ter	5`-CAT CCC GGG CGC TTG CTT TAC CTC GGC AG-3`
BACK Fl_N-ter	5'-GCC AAT TGC TCC TGG TCC GAA AGG TTC ACT TCC TCC-3`

PCR2 primers

FOR Fl_C-ter	5`-TAT TAC AAT TGA TGG CGC AGA TGG CCC GAT AAC GAT-3`
BACK Fl_C-ter	5`- GGG CCG GCA ATA AGC GAA CAT TCT TCC TAC AGA ATT AGA AT-3`

PCR3-6 primers

FOR Flag_N-ter	5`-CCA TGG GCA AAG GGC GGC ATG ATA-3`
BACK Flag_C-ter	5`-CGA GAT CTA TAA GCG AAC ATT CTT CC-3`
BACK Flag_Trunc C-ter	5`-CGA GAT CTA CGC GAG CTA GGA TAA TAA G-3`
BACK Flag_c-mycC-ter	5`-AGG CGG CCG CAT AAG CGA ACA TTC TTC C-3`
BACK Flag_Trunc c-myc_C-ter	5`-ATG CGG CCG CAC GCG AGC TAG GAT AAT AAG-3`

PCR primers for cloning of FLAG N-terminus and FLAG C-terminus into bacterial expression system

PCR1 primers
FOR Flag C-Ter 5`-CCG AAT TCT CGC GTG TTC CCG ATA TG-3`
BACK Flag C-Ter 5`-CCC TCG AGA TAA GCG AAC ATT CTT CCT AC-3`

PCR2 primers
FOR Flag_N-Ter 5`-TTG AAT TCA TGG GCA AAG GGC GGC ATG ATA CG-3`
BACK Flag_N-ter 5`-TCT CGA GTG CAC CTG AAG GTT GCC CTC CAC -3`

PCR3 primers
FOR MaSp1 5`- CCG AAT TCT CAC GCT TGT CAT CAC C-3`
BACK MaSp1 5`-GGC TCG AGT GAG AAA GCT TGG TTT ATA G-3`

PCR4 primers
FOR MaSp2 5`- CCG AAT TCT CTA GAC TAT CAT CTC C-3`
BACK MaSp2 5`-GTC TCG AGA CCG AGG GTA TTT CC-3`
FOR FlagNt-Trunc29 5’-CACATCGCCTCCGCTGCTTTCGAATTCCG-3’
BACK FlagNtTrunc29 5’-TACAGGTCGTCTTCTTCCGGTGGACTCGAGAC-3’
BACK FlagNtTrunc28 5’-CCCTCGAGTCCTCCACCGGAAGAAGACGAC-3’

Sequencing primers

LMB3 5'-CAG GAA ACA GCT ATG AC -3' http://www.mrc-cpe.cam.ac.uk/~phage/*g1p.html)
PHENseq 5'-CTA TGC GGC CCC ATT CA -3' http://www.mrc-cpe.cam.ac.uk/~phage/*g1p.html)

M13-Forward 5'-TGT AAA ACG ACG GCC AGT-3'

SP6 –Reverse 5`-GTA TTC TAT AGT GTC ACC TAA ATA -3`

4.1.6 Molecular markers

BenchMark™ prestained protein ladder [Invitrogen], Western blot bench marker [Fermentas], Western blot positive control scFv purified *via* affinity matrix from bacteria [Phytoantibody group, IPK] DNA Smart ladder [Eurogentec]

4.1.7 Enzymes

Alkaline phosphatase [Amersham], Klenow enzyme [USB], restriction enzymes [Invitrogen, Roche, Fermentas], RNase A [Qiagen], Taq polymerase [Roche], Pwo super yield polymerase [Roche], T4-DNA ligase [Invitrogen], Lysozym (Sigma)

4.1.8 Antibodies and antibody conjugates

Anti c-myc antibodies 9E10 (Munroe and Pelham 1986), anti mouse alkaline phosphatase conjugate [Sigma], anti rabbit IgG alkaline phosphatase conjugate [Amersham Pharmacia Biotech], anti mouse IgG horseradish peroxidise linked whole molecule (IgG-HRP) [Amersham Pharmacia Biotech, D], anti rabbit IgG horseradish peroxidise linked whole molecule (IgG-HRP) [Amersham Pharmacia Biotech, D], anti histidine tag antibody [Sigma]

4.1.9 Antibiotics and plant growth regulators

Ampicilin sodium (Amp) [Duchefa, NL], carbenicillin disodium (Cb) [Duchefa, NL], cefotaxime sodium (Cla) [Duchefa, NL], kanamycin monosulfate (Km) [Duchefa, NL], Rifampicin disodium [Duchefa], NAA (α-naphthaleneacetic acid) [Sigma], BA (6-benzyl-amino-purine) [Sigma].

4.1.10 Kits

ECL Western blotting analysis system [Amersham], QIAEX II gel extraction kit [QIAGEN], QIAGEN plasmid kit [QIAGEN], PGEM-T Vector System I (T/A cloning kit) [Promega], mini-dialysis kit [Amersham], Ni-NTA agarose [Qiagen], protein A agarose [Roche],

4.1.11 Chemical reagents

Bacto®-agar [Difco, USA], bacto®-peptone [Difco, USA], yeast extract [Difco, USA], DEPC [Sigma, D], bovine serum albumin [Sigma, D], p-nitrophenylphosphate (p-NPP) [Sigma, D], bromphenolblue [Sigma, D], ethidium bromide [Roth, D], 37% formaldehyde [Roth, D], formamide [Roth, D], glycine [Roth, D], glycerol [Roth, D] Rotiphorese® Gel30 (Acrylamid/Bisacrylamid), Roth, D], phenol/chloroform [Roth, D], tris [Roth, D], sodiumchloride [Roth, D], methanol [Roth, D], isopropanol [Roth, D], acetic acid [Roth, D], IPTG [Boehringer/Mannheim, D], marvel dried skimmed milk fat free [Marvel, GB], MES [AppliChem, D], murashige-skoog medium basal salt mixture [Duchefa, NL], NBT/BCIP stable mix [GibcoBRL, D], polyethylenglycol 6000 (PEG 6000) [Merck, D], sucrose [Duchefa, NL], sarcosyl, TEMED [Gibco BRL, D], EDTA [Gibco BRL, D], triton X-100 [Serva, D], tween 20 [Serva,D], 2-mercaptoethanol

[Serva], bovine serum albumin (BSA) [Serva], ammonium sulphate [USB], Bio-Rad protein assay [Bio-Rad, D], gelcode® blue stain reagent [Perbio, D], acetosyringone.

4.1.12 Special laboratory materials

Microtitre plates Falcon® 3915 ProBind [Becton Dickinson and company, USA], Parafilm® – laboratory film [American National Can™, USA], ProbeQuant G-50 Micro columns [Amersham, D], Protran® nitrocellulose transfer membrane [Schleicher & Schuell, D], Filter paper Whatman no. 4 [Whatman®, USA], Syringe filters 0.2 µm Nalgene® [Schütt Labortechnik, D], [Whatman®, USA], X-ray film [Retina, D], mortars and pestles, ECL™ Western Blotting Detection Reagents, [GE Healthcare RPN2109]

4.1.13 Special laboratory equipment

Benchtop microcentrifuge Biofuge fresco [Heraeus, D], centrifuge Varifuge 3.0R [Heraeus, D], centrifuge evolution RC [Sorvall], ELISA reader Dynatech MR 7000 [Dynex, D] GeneAmp® PCR system 9700 [PE Applied Biosystems, USA], Gene-Pulser™ [Bio-Rad, D], spectrophotometer Spectronic Genesys 5 [Milton Roy, USA], transfer electrophoresis unit [Schütt Labortechnik, D], UV transilluminator [Intas, D], ultrasonic device Vibracell [Sonics&Materials], vacuum centrifuge concentrator 5301 [Eppendorf, D], gel documentation system with software [Intas], pH-surface electrode [Sartorius], electrophoresis power supply Consort 132, [Turnhout, Belgium], Mini Protean® (mini-gel electrophoresis system for vertical PAAGE) [Bio-Rad, D], Mini Trans-Blot Cell® (mini-blot-apparatus) [Bio-Rad, D], Laminar flow/biohazard hood [Nuare, USA], horizontal orbital shaker "The belly dancer" [Stovall life science, USA], Vortex mixer [Bender, Zurich], UHP (ultrahigh pure water system, Membrane Pure, Bodenheim.)

4.1.14 Buffers

In all solutions, UHP (ultrahigh pure water system)

Buffers for cloning:

TFB1 solution: 100 mM RbCl,
50 mM MnCl2,
30 mM KAc,
10 mM $CaCl_2$,
15% glycerol, pH 8.0
Sterilised by filtering

TFB2 solution: 10 mM MOPS,
75 mM $CaCl_2$,
10 mM $RbCl_2$,

15% glycerol
Sterilised by filtering

Lysis buffer: 50 mM glucose,
10 mM EDTA,
25 mM Tris-HCl, pH 8.0.

Lysozyme: 4 mg/ml lysozyme.

P2 buffer : 0.2 M NaOH, 1% SDS.

P3 buffer : 3 M Sodium acetate, pH 5.2.

TAE buffer for agarose gel:
40 mM Tris-acetate,
1 mM EDTA, pH 8.0

DNA- loading sample buffer:
0.2 M EDTA,
30% Glycerin,
0.1% Bromphenolblue

TE (Tris/EDTA): 10 mM Tris-HCl, pH 8.0,
1 mM EDTA

Buffers for PAGE and Western Blot:

Stacking gel buffer: 0.5 M Tris, pH 6.8,
0.4% SDS.

Separating gel buffer: 1.5 M Tris, pH 8.8,
0.4% SDS.

SDS- gel running buffer :
0.025 M Tris,
0.25 M Glycin,
0.1% SDS, pH 8.3.

2xSDS-sample buffer: 100 mM Tris, pH 6.8,
20% glycerine,
4% SDS,
10% β-mercaptoethanol, 0.2% bromphenolblue

Transfer buffer: 10% Methanol,
gel running buffer

4xMarvel buffer: 80 mM Tris-HCl pH 7.8,
720 mM NaCl.

Buffers for ELISA:

PBS: 8 mM Na_2HPO_4,
2 mM KH_2PO4,
150 mM NaCl, pH 7.2.

Alkaline phosphatise (AP)

Substrate buffer: 97 ml Diethanolamin,
0.1 g $MgCl_2$,
0.2 g NaN_3, pH 9.7

Buffers for recombinant protein purification from *Escherichia coli* and plant crude extract:

Suspension buffer: 20 mM TRIS-HCl
1 mM EDTA, pH 8.0

Starting buffer: 100 mM $Na_2HPO_4.12H_2O$,
100 mM $NaH_2PO_4.H_2O$, pH 7.0.

Binding buffer: 20 mM Tris-HCl,
300 mM NaCl,
10 mM Imidazole, pH 7.9

Washing buffer: 30 mM Imidazole,
300 mM NaCl,
20 mM Tris-HCl, pH 7.9

Elution buffer: 250 mM Imidazole,
300 mM NaCl,
20 mM Tris-HCl, pH 7.9

Elution buffer (EB): 0.2 M Glycin pH 2.5.

Neutralisation solution:
1 M Tris-HCl pH 8.0.

4.1.15 Media

All media are calculated for a volume 1 L and were sterilised by autoclaving for 20 min at 120°C. Antibiotics, hormones or other components were added from filter sterilised (0.2 µm pore size) stock solutions after autoclaving the media.

LB medium: 5 g yeast extract,
10 g bacto-peptone,
10 g NaCl, pH 7.0
for a solidified medium 15 g bacto-agar was added per 1 L liquid medium,
sterilised by autoclaving

Murashige-Skoog medium:
4.49 g Murashige-Skoog medium basal salt mixture (MS),
0.5 g MES,
30 g Sucrose, pH 5.7,
sterilised by autoclaving

Callus induction medium:
0.8-1.0% agar was added for solidified medium. 0.1 mL NAA (α-naphthaleneacetic acid) and 1.0 mL BAP (6-benzyl-amino-purine) were added from 1 mg/mL stock solutions. To make stock solutions, 100 mg powder was dissolved in a small volume (1-2 mL) of 1N NaOH, made up to 100 mL with distilled water and filter sterilised (0.2 µM pore size), pH 5.7

MG3 medium- Murashige-Skoog-Glucose medium:
16 g Glucose

2 mg Zeatin ribose,
20 µg NAA,
20 µg GA3,
250 mg Cefataxim,
50 mg Kanamycin,
(9 g agar was added for solidified medium)

SOC medium: 5 g yeast extract,
20 g tryptone,
0.58 g NaCl,
0.186 g KCl.,
1 mL of 2 M glucose (to the final concentration 20 mM) was added after autoclaving, pH 7.0.

2xTY medium: 10 g yeast extract,
16 g tryptone,
5 g NaCl, pH 7.0,
10 g agar for solidified medium,
sterilised by autoclaving

YEB medium: 5 g beef extract,
1 g yeast extract,
5 g peptone,
5 g sucrose,
5 ml of 2 M $MgSO_4$, pH 7.2,
sterilised by autoclaving

YEB medium appropriated for *vir* gene induction:
YEB medium,
10 mM 2-(N-morpholino) ethanesulfonic acid (MES), pH 5.6,
antibiotics,
20 µM acetosyringone.

MMA medium: MS salts,
10 mM MES,
20 g/L sucrose, pH 5.6,
200 µM acetosyringone.

2xInfiltration medium: 100 g sucrose,
3.6 g glucose,
8.6 g MS

4.1.16 Software and special equipment

DNA and protein sequence data were processed using the programme package Vector NTI suite 6.0 and Vector NTI advance 10 programme (Invitrogen), Scanner “HP Scanjet 6300C” (Hewlett Packard)

4.2 Methods

4.2.1 Cloning and sequencing

All standard genetic engineering techniques (propagation of bacterial strains, agarose gel electrophoresis, restriction and ligation) were performed according to Sambrook et al. (1989). Plasmid DNA extraction and purification was done using the QIAGEN plasmid kit. Plasmid DNA was extracted from agarose gels using the QIAEX Gel extraction kit. For sequencing purposes, plasmid DNA was purified from 3 mL cultures of bacteria with the QIAGEN kit. DNA sequence determination, according to Sanger et al. (1977), was done with universal primers by MWG company (http://www.mwg-biotech.com).

4.2.2 PCR

The mixture was made in a 0.5 mL "safe-lock" microcentrifuge tube:

Primer 1 (10 pmol)	1 µL
Primer 2 (10 pmol)	1 µL
5 ng template DNA	5 µL
10 mM dNTPs	2 µL
10X PCR-buffer	5 µL
Enzyme Taq Polymerase	1 µL
H_2O	35 µL
Total volume	50 µL

All pipeting was done on ice. PCR reaction was carried out using the following parameters: 30 cycles of 95 ºC-50s, 50 ºC-30s, and 72 ºC-40s
The amplified PCR fragment was separated in 1% agarose gel in TAE.

4.2.3 Transformation of *Escherichia coli*

4.2.3.1 Heat shock competent cells

5 mL LB medium was inoculated with a single colony of XL1 Blue cells and incubated overnight at 37 ºC. 1 mL of the overnight culture was used for the inoculation of 100 mL LB medium. The culture was incubated at 37ºC by shaking until the optical density at 600 nm reached 0.5. Cells were centrifuged for 10 min at 4ºC (4500 rpm) and the pellet was resuspended in 30 mL ice cold solution TFB1 and let on ice for 1 h. Cells were centrifuged for 10 min at 4ºC (4500 rpm) and the pellet was resuspended in 4 mL ice cold solution TFB2. 100 µL aliquots of heat shock-competent cells were frozen in liquid nitrogen and stored at -80ºC.

4.2.3.2 Heat shock transformation

100 µL of heat shock-competent XL1 Blue cells were incubated for 15 min on ice with 1 µL of plasmid DNA and afterwards heated for 90 s at 42ºC. Then mix of plasmid DNA with competent cells were placed on ice for 3 min. Cells were resuspended in 500 µL of LB medium, incubated for 10 min at 37ºC and plated on a solidified LB medium with appropriate antibiotics.

4.2.4 Total DNA isolation from *Agrobacterium tumefaciens*

A single bacterial colony was taken for inoculation of 2 mL LB medium containing the appropriate antibiotics and left to culture for 36-48 h at 28ºC, under constant rotation. The cells were collected by centrifuging 1.5 mL culture at 13000 rpm/5min and the supernatant was completely removed. The pellet was resuspended in 100 µL cold buffer P1 (Qiagen) followed by addition of 200 µL buffer P2 (Qiagen), carefully mixed by inverting microcentrifuge tubes several times and incubated at RT for 5 min. 150 µL of buffer P3 (Qiagen) were added to mixture, vortexed for 10 s and incubated 10 min on ice. After further centrifugation (13000 rpm/20 min), the upper DNA-containing aqueous phase was transfered into fresh tubes. 0.7 v/v of isopropanol was added; the mixture was inverted several times and incubated at RT for 10 min. After further centrifugation (13000 rpm/30 min), the supernatant was removed and the pellet washed with 2.5 volumes of 70% ethanol by carefully inverting the microcentrifuge tubes and recentrifuged. After removal of the supernatant, the DNA pellet was dried and then dissolved in 30 µL water.
Enzyme restriction was carried out using 10 µL DNA solution, 2 µL HindIII enzyme (15 U/µL), the particular reaction buffer and made up to a volume of 20 µL. The reaction mixture was incubated at the appropriate temperature for 2-3 h. DNA fragments were electrophoretically separated on a 1% agarose gel.

4.2.5 Transient expression of recombinant proteins in tobacco leaves

Transformation of *Agrobacterium tumefaciens* by high-voltage electroporation
The transformation of *Agrobacterium* by electroporation is a quick and simple method for the introduction of plasmid DNA into competent cells.

4.2.6 Preparation of competent cells

10 mL SOC medium containing appropriate antibiotics was inoculated with a single colony from a plate of *Agrobacterium tumefaciens* and incubated in a shaker (150-200 rpm) at 28ºC for 2 days 100 mL of SOC medium containing kanamycin was inoculated with 0.5 mL of pre-culture and incubated as above. Cells were harvested by centrifugation, in a Sorvall SS34 rotor, 4300 g, 15 min, at 4ºC. The pellet was resuspended in 100 mL sterile, distilled water and centrifuged. The cells were twice more washed with 50 mL ice cold water and finally in 10 mL 10% (v/v) glycerol. After

centrifugation, the cells were resuspended in 500 µL 10% (v/v) glycerol and aliquoted into sterile microcentrifuge tubes, freezed in liquid nitrogen and stored at -80ºC.

4.2.6.1 Electroporation

An aliquot of cells was thawed on ice. 10-100 ng of plasmid DNA was mixed and transferred to an ice-cold electroporation cuvet. Electroporation was perfomed with a Bio-Rad Gene Pulser ™ at 25 µF, 2.5V, 200Ω. The cuvet was rinsed out with 1 mL SOC medium and was transferred to a sterile microcentrifuge tube. The cells were incubated by shaking gently at 28ºC for 2 h. 100 µL aliquot was spread onto selective SOC plates and incubated at 28ºC for 48 h until colonies appeared.

4.2.6.2 Infiltration

5 mL YEB-medium containing rifampicin, carbenicilin, kanamycin was inoculated with single colony from *Agrobacterium tumefaciens* plate and incubated in shaker (180 rpm) at 28ºC.
After 24 h incubation of medium with *Agrobacterium*, acetosyringone was added to final concentration 20 µM, glucose to 10 mM, 2-(N-morpholino) ethanesulfonic acid (MES) pH 5.6 to 10 mM and cells were further grown.
On third day OD of culture was determined at 600 nm wave length and adjusted to 1.0 with 2x infiltration medium and water. Acetosyringone was added to final concentration 200 µM and then used for infiltrations.
Agro-suspension was injected with 1mL syringe into leaves and plants were incubated for 3 days (23ºC, 16 h light).
Leaf discs were collected into microcentrifuge tubes, protein extracts were made and analysed by Western blot.

4.2.7 Protein extraction from plant material

Two leaf discs were grinded using a drilling device in 100 µL of SDS-sample buffer with mercaptoethanol. The homogenate was boiled for 10 min and then centrifuged for 20 min at 4ºC, after which the supernatant was transferred to a fresh tube. The total protein content of this fraction was determined according to the method of Bradford (1976). Here, 1 µL extract was mixed with 799 µL water and 200 µL protein assay reagent and absorbance measured at 595 nm. Sample protein content was estimated according to BSA standards.

4.2.8 Western blot analysis

5-20 µg of total soluble protein extract were separated by PAGE under denaturing conditions, using a 10% acrylamide gel. Both the gel and electrophoresis buffer contained SDS. Prestained protein standards were used as size markers. The transfer of proteins to a

nitrocellulose membrane was carried out electrophoretically in SDS-containing buffer (250 mA, 18V, 8-12 h). Unspecific sites on the membrane were blocked by incubating it in 5% Marvel (5% dry milk in Marvel buffer) for 2 h with constant gentle rotation. A 1:50 dilution of the hybridoma cell culture supernatant, containing the 9E10 anti c-myc antibody (Munroe and Pelham 1986), in 0.5% Marvel was made and the filter incubated in this for 2 h at RT and afterwards the blots were washed three times in 0.5% Marvel for 10 min before application of the secondary antibody, anti mouse IgG peroxidase conjugate, diluted 1:2000 in 0.5% Marvel for 1 h/RT. The filter was thoroughly washed twice in 0.5% Marvel, twice in 1x Marvel buffer and twice in PBS buffer. Detection of the fusion protein on the filter was with chemiluminescence, using the ECL system. A substrate was applied for 1 min, the blot immediately wrapped in cellophane and incubated with X-ray film for 1-5 min. Expression levels were estimated by comparison of the intensity of signals from protein samples with those of defined quantities of purified scFv used as a positive control.

4.2.9 Leaf disc transformation of potato

Plant transformation was carried out using *Agrobacterium tumefaciens*-mediated gene transfer (Zambriski et al. 1983). Potato plants grown under sterile conditions as *in vitro* cultures were used for leaf disc transformation. Young, though full sized leaves of potato were cut into square pieces (about 10 mm in width) using a sterile scalpel blade. A 50 µL of fresh overnight *Agrobacterium* culture was added in 10 mL liquid MS medium, the explants dipped into this and then placed on low light conditions for 2 days at 24ºC. The leaf explants were then transferred to plates of MG medium containing 5 mg/L NAA and 0.1 mg/L BAP, though containing kanamycin (50 mg/L) and cefotaxime (250 mg/L). After 1-2 week later the explants were placed onto the fresh MG3-medium to allow regeneration of shoots. 3-4 weeks later, developing shoots were excised from the callus and placed on MS including antibiotics but without growth regulators. Those shoots which developed roots were then planted in small pots of compost and kept in a propagator for 2 weeks before acclimatisation and repotting.

4.2.9.1 Plant propagation *in vitro*

For multiplication, the stems were cut in nodal sections, each including an auxiliary bud. The leaf from each section was trimmed off, leaving a short petiole stump. The sections bases were placed down onto the selection medium. Plants were sub cultured every 4 to 5 weeks.

4.2.10 Purification of recombinant spider silk-like proteins from plant material

4.2.10.1 Purification of Fc-100XELP

5 g of the frozen tobacco leaf was grinded in liquid nitrogen and 3 volumes of 0,15 M PBS containing 0.1% of Triton X100 (PBS-T) were added. Approximately 0.1 mg of a frozen tobacco leaf were extracted in SDS-sample buffer and used as a control of the Fc-100xELP expression. The samples were centrifuged at 4°C for 10 min (13000 rpm) and the supernatant centrifuged again in 2 mL microcentrifuge tubes at 4°C for 10 min (13000 rpm). Sodium chloride was added to the supernatant (to a final concentration of 2,75 M PBS) and incubated in a waterbath for 20 min at 60ºC. The sample was centrifuged 30 min at 40ºC. The pellet was resuspended in 1.5 mL of PBS buffer and incubated at RT by gentle shaking on the wheel. The sample was centrifuged two times at 4ºC for 10 min (13000 rpm) and the pellet was resuspended in 2 mL PBS-Triton buffer. The application of the supernatant onto the column was repeated two times. 0.5 mL of the crude plant extract was used as a control of the extraction. 0.5 mL of the solution after purifying *via* the column was used as a control of the binding effectivity. The protein G column was washed with 30 mL of PBS buffer. Fc-100xELP recombinant proteins were eluted with 16.5 mL of 0.2 M glycine pH 2.5, into 2 mL microcentrifuge tubes, each containing 400 µL of 1 M Tris pH 8.0 for neutralisation. The column was washed with 100 mL of PBS buffer followed by the application of 10 mL in PBS buffer containing Na-Azid. A dot blot analysis was performed to detect recombinant proteins in the fractions and in through flows. Fractions containing recombinant protein Fc-100xELP were combined and filled in dialysis tubing. Dialysis was performed first at RT against PBS carefully stirred on magnetic mixer. Then the samples were concentrated in PEG 6000 at RT for 1.5 h. Finally, protein was dialysed against PBS overnight at RT. The protein was lyophilised and weighed.

4.2.10.2 Purification of SO1-100XELP by heat treatment and inverse transition cycling according to (Scheller 2004)

Spider silk-elastin fusion proteins were extracted with 2 L of raw extract buffer from 1 kg fresh or frozen leaf material or tubers in a standard waring blender. Subsequently, the extract was heated at 95ºC for 60 min and was cleared by centrifugation at 8000 g for 30 min. The supernatant was incubated in a water bath at 95ºC for 60 min. The precipitate was removed by centrifugation at 8000 g for 30 min at 4ºC. The supernatant was transferred into fresh glass and sodium chloride added to final concentration 2 M/L. To precipitate the elastin the solution was incubated in a water bath at 60ºC for 60 min. The precipitate was removed by centrifugation at 8000 g for 30 min at 40ºC. The precipitate was dissolved in 80 mL raw extract buffer (PBS without sodium chloride) for 15 min on a shaker at 20ºC. Insoluble material was removed by centrifugation at 8000 g for 15 min at 20ºC. The supernatant was dialysed against 5 L PBS-buffer. The protein solution was sterilised by filtration. Purified spider silk-elastin proteins could be stored at -20ºC. In each step from supernatants samples were made and tested by Western blot for the presence of synthetic spider silk-elastin proteins.

4.2.11 Expresssion of recombinant proteins in *Escherichia coli*

Escherichia coli strain BL21 (DE3) was transformed with plasmid pET23a containing FLAG CT, MaSp1 CT, MaSp2 CT and FLAG NT variants and recombinant proteins were expressed.
Proteins were produced in the *Escherichia coli* strain BL21 (DE3). Cells were grown at 37ºC in LB-medium supplemented with ampicillin (100 µg/mL). Pre-culture of cells grown to a log-phase stored overnight at 4ºC. The next day 0.5 L of fresh medium was inoculated with 10 mL of pre-culture, flasks were vigorously shaken at 37ºC. When the culture reached an optical density at 600 nm of 0.4-0.5, isopropyl β-d-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM and cultivation was continued for 3 h at 37ºC. Cells were centrifuged, washed, resuspended in 15-20 mL binding buffer. At this step cells could be frozen in liquid nitrogen and stored at -80ºC.

4.2.11.1 Cell lysis

Cells were thawed at RT for 15' and then kept on ice. Suspensions of cells were disrupted on ice using a sonicator equipped with a thick tip (13 mm in diameter). Sonication was carried out on ice with 5s on/ 5s off cycles until a clear lysate was evident. Sonication procedure was continued 500 s at 12-16 Watt. The cell lysate were centrifuged for 30 min at 12000 rpm, at 4ºC, using Sorvall SS-34 rotor. Supernatant were transferred into fresh falcon tube, and kept on ice ready for further use in purification. In case when target protein is insoluble the pellet was used for further purification. Therefore the pellet was resuspended in equal volume of binding buffer containing 6 M Guanidine hydrochloride as solubilising agent for insoluble proteins. The protein concentration was determined and 20 µg soluble and insoluble protein extracts were separated by PAGE, using 15% gel, after boiling in the same volume of SDS sample extraction buffer for 10 min.

4.2.11.2 Purification of recombinant proteins using Ni-NTA agarose under native conditions

Columns for purification were rinsed with 3 mL of ddH_2O. Ni-NTA agarose was well mixed and 1.8 mL were transferred into 2mL microcentrifuge tubes, centrifuged for 1min at 800 rpm and the supernatant was carefully removed. 1 mL of ddH_2O was added to agarose, mixed and again centrifuged. The procedure was repeated one time more and after the last washing step the slurry was poured into the column and the tip of the column was capped, allowing the resin to settle one hour at RT. The 5 mL ddH_2O was allowed to flow through by uncapping the bottom outlet. The column was equilibrated by adding of 10 mL binding buffer. The supernatant containing recombinant protein was applied to a Ni-NTA agarose column to bind His-tagged proteins Adsorption, washing and elution was performed as described by the manufacturer (Qiagen, Novagen). Glycerol was added to the eluted proteins to a final concentration of 50% to prevent precipitation. Proteins were stored at -20ºC.

Samples were taken after each step of purification procedure and SDS-samples were prepared and subjected to PAGE analysis. Therefore the protein concentration was determined and 20 µg soluble and insoluble protein extracts were separated by SDS-PAGE, using a15% gel, after boiling in the same volume of SDS sample extraction buffer for 10 min.
Purification procedure with insoluble proteins was performed similar way with buffers containing 6 M Guanidine hydrochloride.

4.2.12 Production of soluble antibodies (scFv)

5 mL of 2xTY + Ampicillin + 0.1% Glucose were infected with single colonies of *Escherichia coli* cells (HB2151) containing pIT2 plasmids with gene sequences encoding scFv's and culture were grown at 37ºC until OD 0.7-0.9.
Protein expression was induced with IPTG to final concentration of 1 mM and incubation continued with shaking (300 rpm) at 30ºC overnight. Cells were pelleted by centrifugation at 4500 rpm 10 min at 4ºC. Supernatant were transferred to fresh falcon tube, and stored at -20ºC or directly used for purification of scFv using Ni-NTA affinity chromatography. Samples were taken in microcentrifuge tubes and used as primary antibody to perform scFv ELISA.

4.2.13 Indirect scFv ELISA

Antigens were appropriately diluted in PBS bufferto (FNT 1 µg/mL, FCT 10 µg/mL). A microtiter plate was coated with antigens by adding 100 µL of the protein solution per well. Plates were incubated for 18-20h at RT. Unoccupied sites were blocked with a blocking agent (200 µL/well) with 3% BSA in PBS buffer. Plates were stored at RT for 2 h. Samples of antibodies (scFv's) were added (100 µL/well) in dilution 1:1 with BSA-PBS. Wells of plates were washed 5x with PBS-Tween (0.1% Tween 20). Anti c-myc antibodies (in dilution 1:50) in 1% BSA-PBS were added (100 µL/well) and incubated 1 h at RT. Washing of wells were performed again 3x with PBS-T buffer. Enzyme-antibody (anti mouse alkaline phophatase conjugate) was added diluted (1:2000) in 1% BSA-PBS buffer (100 µL/well) and incubated for 1 h. Washing step was repeated again 3x times with PBS-T. Substrate p-nitrophenilphosphate (pNPP) dissolved in substrate buffer (SB) were added to the wells (100 µL/well) and incubated 1 h at 37ºC. OD was measured at 405 nm wavelength using an ELISA reader. The results were analysed by Microsoft Excel 2000.

4.2.14 Immunisation of rabbits with recombinantly produced antigen

Rabbits were immunised with 1mg recombinant protein and complete Freud adjuvant subcutaneously; boosted after 1 week and 2 weeks with each 0.5 mg recombinant protein and incomplete Freud adjuvant. After checking the titre by ELISA and Western blot a further immunisation with 0.5 mg and incomplete Freud adjuvant was performed if necessary. After a sufficient titre was achieved, the rabbits were bled and serum was precipitated by 50% saturated ammonium sulphate.

4.2.15 Ammonium sulphate precipitation of rabbit serum

To a 30 mL of rabbit serum were added by dropping equal volume of saturated ammonium sulphate solution and stirred on magnetic mixer for 2 h at 4ºC. The solution was centrifuged at 4000 rpm, for 15 min at 4ºC. The pellet was resolubilised in 30 mL PBS, pH 7.6. The antibody solution were dialysed 3x times against PBS pH 7.6 at 4ºC. The antibody solution was aliquoted and stored at -20ºC.

5.1 Summary

Until now, various experiments to provide sufficient amount of feedstock of recombinant spider silk proteins for structural studies or for spinning experiments were performed. For these purposes native or synthetically constructed consensus repeats of dragline silk proteins MaSp1 and 2 were employed. As heterologous production systems different hosts as bacteria, yeasts, animal cells, insect cells and plants were tested. Hovewer, flagelliform silk of *Nephila clavipes* could serve as potential target for structural studies of spider silk proteins in terms of elasticity. Since 1998 partial cDNA clones encoding N-terminus, C-terminus and repetitive parts of FLAG have been described. It would be of great benefit to target products of these gene sequences to structural investigations.
The main interest of our study was to design a *FLAG* gene with its non-repetitive N- and C-terminal parts and with a repetitive part displaying its main modular structure. In order to achieve expression of FLAG protein in the ER of plant cells, cDNA clones encoding distinct parts of flagelliform silk protein derived from *Nephila clavipes* were combined into one continuous reading frame.
The special aim of the design of *FLAG* gene variants was to avoid any tag to be used for immunodetection. Therefore, non-repetitive N- and C-termini were bacterially expressed and employed as antigens for raising antibodies. The powerful tool of recombinant antibody engineering allowed us to raise specific antibodies, which were successfully applied in immunodetection of FLAG protein from plant crude extract. Cloned variants of the *FLAG* gene were successfully tested in a transient expression system in protoplasts for expression of the FLAG protein.
Genes encoding the *FLAG* variants were subsequently cloned into binary vectors for ubiquitous expression in plants. The constructs were tested for functionality in transient expression system by agroinfiltration in tobacco leaves. The Western blot analysis showed successful expression of all four constructs in a binary vector. *Agrobacterium*-mediated leaf disc transformations were performed in tobacco and stable transformants are produced. For each construct a number of high expressor plants were selected. To establish production lines, best producer plants were estimated and further propagated.

In order to evaluate potato tubers as biofactories for big scale production of spider silk proteins plant expression vectors for ubiquitous spider silk-ELP fusion proteins accumulation in the ER were used for transformation into starch potato. Transgenic potato plants were generated by *Agrobacterium*-mediated transformation. The expression of the spider silk-ELP fusion proteins was detected by Western blot in leaves and tubers of transgenic tobacco and potato plants. In greenhouse experiments the diameter of tubers, the morphology and the tuber yield was measured in different lines of the constructs MaSp1-100xELP and MaSp2-100xELP and SO1-100xELP. The diameter of tubers was observed to differ from line to line. The relation between tuber and leaf expression was different in the individual lines. The yield of tubers from greenhouse plants was dependent on growth conditions and seasonable differences. The transgenic potato lines were described and used in field trial. The importance of obtaining properly folded high molecular mass soluble fusion proteins becomes apparent when transforming the corresponding genes into plants for molecular farming. Therefore, studies about

increasing of size of protein molecules post-translationally were one of the subjects of this study.
Cloning of Fc-100xELP gene into pRTRA 7/3 vector with subsequent cloning into binary vector allowed the production of Fc-100xELP fusion proteins in tobacco and potato plants. The use of *Agrobacterium*-mediated gene transfer enabled production of stable transformants. The Fc-100xELP fusion protein was shown to form dimers under non-reducing conditions in Western blot analysis. Fc-100xELP fusion proteins were purified using the protein A affinity chromatography method. The purified proteins are under investigation now according their structural and physical properties.

5.2 Zusammenfassung

In letzter Zeit wurde über verschiedene Versuche berichtet, ausreichende Mengen rekombinanter Spinnenseideproteine für Strukturuntersuchungen und Verspinnungsexperimente bereitzustellen. Für diese Zwecke wurden native oder synthetisch konstruierte Konsensussequenzen der Spinnenseidenproteine MaSp1 und MaSp2 benutzt. Verschiedene heterologe Produktionssysteme wie Bakterien, Hefen, tierische Zellen, Insektenzellen und Pflanzen wurden getestet. Im Gegensatz zu MaSp1 und MaSp2 eignet sich Flagelliform-Seide der Spinne *Nephila clavipes* als potentielles Objekt für strukturelle Studien an Spinnenseidenproteinen in Hinblick auf Elastizität von entsprechenden Materialien. Seit 1998 sind partielle cDNA-Sequenzen, die den N-Terminus, den C-Terminus und repetitive Teile des Flagelliformseidenproteins FLAG kodieren, identifiziert wurden. Die Genprodukte dieser Sequenzen strukturellen Untersuchungen zugänglich zu machen, wäre von großem wissenschaftlichen als auch von praktischem Interesse. Eine wichtige Aufgabe der vorliegenden Studie war daher, ein *FLAG*-Gen zu konstruieren, das sowohl N- und C-terminale nichtrepetitive Teile als auch einen repetitiven Teil enthält und daher im Wesentlichen die modulare Struktur des nativen *FLAG*-Genes widerspiegelt. Um das FLAG-Protein im ER von Pflanzenzellen zu exprimieren, wurden cDNA-Klone, die bestimmte Teile des FLAG-Proteins kodieren, zu jeweils kontinuierlichen Leserahmen zusammengefügt. Ein Ziel dieser Konstruktionen bestand darin, jegliche spezielle Detektionssequenzen zu vermeiden. Deshalb wurden nichtrepetitive N- und C-terminale Teile bakteriell exprimiert und als Antigene zur Produktion spezifischerAntikörper benutzt. Spezifische rekombinante Antikörper und spezifische konventionelle Antikörper wurden erzeugt, charakterisiert und zum Nachweis rekombinanter FLAG-Proteine aus Pflanzen genutzt. Die konstruierten *FLAG*-Genvarianten wurden erfolgreich durch transiente Expressionsversuche in Protoplasten getestet und anschließend in einen Binärvektor kloniert. Durch Agroinfiltration und anschließende Western Blot-Analysen wurden diese Vektoren ebenfalls erfolgreich getestet und mit Hilfe der Blattscheibenmethode stabil transformierte Tabakpflanzen erzeugt. Für jedes der 4 Konstrukte wurden TG0-Linien mit hoher FLAG-Produktion identifiziert, die für weitere Untersuchungen zur Verfügung stehen.

Eine zweite Aufgabe dieser Arbeit bestand darin, die Stärkekartoffelknolle als Produktionssystem für Spinnenseidenproteine zu testen. Hierzu wurden stabil transformierte transgene Stärkekartoffelpflanzen und trangene Tabakpflanzen erzeugt, die MaSp1-100xELP, MaSp2-100xELP und SO1-100xELP im ER von Blatt- und Knollenzeilen akkumulierten. In Gewächshausexperimenten wurden außer der Transgenexpression der Knollendurchmesser, die Knollenmorphologie und das Gesamtknollengewicht untersucht. Der Knollendurchmesser variierte stark zwischen den transgenen Linien. Das Verhältnis zwischen Blatt- und Knollenexpression war auch innerhalb der Linien sehr unterschiedlich. Das Gesamtknollengewicht von Gewächshauspflanzen wurde von den Wachstumsbedingungen und saisonalen Unterschieden stark beeinflusst. Die so charakterisierten transgenen Stärkekartoffellinien wurden in Freisetzungsversuchen verwendet und charakterisiert.

Für die werkstoffmechanische Charakterisierung ist ein möglichst hohes Molekuargewicht der transgenen Spinneseidenproteine notwendig. Ein weiteres Ziel dieser Studie war, durch posttranslationales Processing das Molekulargewicht von Faserproteinen zu erhöhen. Ein erster Versuch hierzu bestand darin, ein ELP-Immunglobulin-Fc-Fusionsprotein im ER von transgenen Tabakblattzellen zu akkumulieren. Wir konnten zeigen, dass diese Proteine durch Disulfidbrückenbildung wie erwartet Dimere bilden, die unter nichtreduzierenden Bedingungen durch SDS-PAGE und Western blot nachgewiesen wurden. Diese Proteine wurden durch Protein A-Affinitätschromatografie gereinigt und werden hinsichtlich ihrer physikalischen Eigenschaften charakterisiert.

6 References

1. Andersen,S.O. Amino acid composition of spider silks. *Comp. Biochem. Physiol.* **35**, 705-711 (1970).

2. Arcidiacono,S., Mello,C., Kaplan,D., Cheley,S., & Bayley,H. Purification and characterisation of recombinant spider silk expressed in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **49**, 31-38 (1998).

3. Artsaenko,O., Weiler,E.W., Muentz,K., & and Conrad U. Construction and functional characterisation of a single chain Fv antibody binding to the plant hormone abscisic acid. *J. Plant Physiol.* **144**, 427-429 (1994).

4. Artsaenko O. Immunomodulation of ABA activity in transgenic tobacco plants. Thesis/Dissertation. Martin-Luther-Universitaet Halle-Wittenberg, Halle, Germany. 1996

5. Artsaenko,O., Peisker,M., Nieden,U., Fiedler,U., Weiler,E.W., Muentz,K., & and Conrad,U. Expression of a single-chain Fv antibody against abscisic acid creates a wilty phenotype in transgenic tobacco. *The Plant Journal* **8**, 745-750 (1995).

6. Barr,L.A., Fahnestock,S.R., & and Yang,J. Production and purification of recombinant DP1B silk-like protein in plants. *Molecular Breeding* **13**, 345-356 (2004).

7. Beckwitt,R. & Arcidiacono,S. Sequence conservation in the C-terminal region of spider silk proteins (spidroin) from *Nephila clavipes* (Tetragnathidae) and *Araneus bicentenarius* (Aranidae). *The journal of biological chemistry* **269**, 6661-6663 (1994).

8. Bell,A.L. & and Peakall,D.B. Change in fine structure during silk protein production in the ampullate gland of the spider *Araneus diadematus*. *J. Cell Biol.* **42**, 284-295 (1969).

9. Bini,E., Knight,D.P., & Kaplan,D.L. Mapping domain structures in silks from insects and spiders related to protein assembly. *Journal of Molecular Biology* **335**, 27-40 (2004).

10. Bradford,M.M. Rapid and quantitative method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-252 (1976).

11. Candelas,G.C., Ortiz,A., & and Molina,C. The Cylindrical or Tubuliform Glands of *Nephila clavipes*. *The journal of experimental zoology* **237**, 281-285 (1986).

12. Candelas,G.C., Arroyo,G., Carrasco,C., & and Dompenciel,R. Spider silkglands contain a tissue-specific alanine tRNA that accumulates *in vitro* in response to the stimulus for silk protein synthesis. *Developmental biology* **140**, 215-220 (1990).

13. Clackson,T., Hoogenboom,H.R., Griffiths,A.D., & and Winter,G. Making antibody fragments using phage display libraries. *Letters to Nature* **352**, 624-628 (1991).

14. Colgin M.A. & Lewis,R.V. Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like ``spacer regions". *Protein Sci* **7**, 667-672 (1998).

15. Conrad.U. Polymers from plants to develop biodegradable plastics. *TRENDS in Plant Science* **10**, 511-512 (2005).

16. Craig C.L. The evolution of silks spun by arthropods. *Annu. Rev. Ent.* **42**, 231-267 (1997).

17. Craig,C.L. & and Riekel,C. Comparative architecture of silks, fibrous proteins and their encoding genes in insects and spiders. *Comparative Biochemistry and Physiology* **Part B**, 493-507 (2002).

18. Cregg,J.M., Barringer,K.J., Hessler,A.Y., & Madden,K.R. *Pichia pastoris* as a host system for transformations. *Mol. Cell. Biol.* **5**, 3376-3385 (1985).

19. Cunniff,P.M., Fossey,S.A., Auerbach,M.A., & and Song,J.W. in Silk Polymers. eds. D.L.Kaplan, W.W.Adams, B.Farmer, & and C.Viney 185-195 (American Chemical Society, Washington; 1994).

20. Deblaere,R., Bytebier,B., de Greve,H., Deboeck,F., Schell,J., van Montagu,M., & and Leemans,J. Efficient octopine Ti plasmid-derived vectors for *Agrobacterium*-mediated gene transfer to plants. *Nucleic Acids Res.* **13**, 4777-4788 (1985).

21. Denny,M.W. in The Mechanical Properties of Biological Materials: Society for Experimental Biology Symposium, 34. eds. J.F.V.Vincent & J.D.Currey 245-271 (Cambridge University Press, Cambridge; 1976).

22. Doolittle,R.F. & Bork,P. Evolutionary mobile modules in proteins. *Sci. Am.* **269**, 50-56 (1993).

23. Fahnestock,S.R. & Bedzyk,L.A. Production of synthetic spider dragline silk protein in *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* **47**, 33-39 (1997).

24. Fahnestock,S.R. & Irwin,S.L. Synthetic spider dragline silk proteins and their production in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **47**, 23-32 (1997).

25. Fiedler,U.a.C.U. High-level production and long-term storage of engineered antibodies in transgenic tobacco seeds. *Biotechnology* **13**, 1090-1093 (1995).

26. Fischer,R. & Emans,N. Molecular farming of pharmaceutical proteins. *Transgenic Res.* **9**, 279-299 (2000).

27. Fischer,R., Stoger,E., Schillberg,S., & Christou,P.a.T.R.M. Plant-based production of biopharmaceuticals. *Current Opininon in Plant Biology* 152-158 (2004).

28. Folk,E.J. Mechanism and basis for substrate specificity of transglutaminase-catalysed epsilon -(gamma-glutamyl)lysine bond formation. *Adv. Enzymol. Relat. Areas Mol. Biol.* **54**, 1-56 (1983).

29. Foo,C.W.F., Bini,E., Hensman,J., Knight,D.P., Lewis,R.V., & Kaplan,D.L. Role of pH and charge on silk protein assembly in insects and spiders. *Appl. Phys.* **A 82**, 223-233 (2006).

30. Foo,C.W.P. & Kaplan,D.L. Genetic engineering of fibrous proteins: spider dragline silk and collagen. *Advanced Drug Delivery Reviews* **54**, 1131-1143 (2002).

31. Gahrtz,M. & and Conrad,U. ed. In press 1-31 (2007).

32. Gatesy,J., Hayashi,C., Motriuk,D., Woods,J., & Lewis,R. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science* **291**, 2603-2605 (2001).

33. Gosline,J.M. & Denny,M.W.a.D.M.E. Spider silk as rubber. *Nature* **309**, 551-552 (1984).

34. Gosline,J.M., DeMont,M.E., & and Denny,M. The structure and properties of spider silk. *Endeavour* **10**, 37-43 (1986).

35. Gosline,J.M., Guerette,P.A., & Ortlepp,C.S.a.S.K.N. The mechanical design of spider silks: from fibroin sequence to mechanical function. *The journal of experimental biology* **202**, 3295-3303 (1999).

36. Guerette,P.A., Ginzinger,D.G., Weber,B.H.F., & Gosline,J.M. Silk properties determined by gland-specific expression of a spider fibroin gene family. *Science* **272**, 112-115 (1996).

37. Hayashi,C. & and Lewis,M. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. *J. Mol. Biol.* **275**, 773-784 (1998).

38. Hayashi,C., Shipley,N.H., & Lewis,R.V. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. *International journal of Biological Macromolecules* **24**, 271-275 (1999).

39. Hayashi,C. & andLewis,R.V. Molecular architecture and evolution of a modular spider silk protein gene. *Science* **287**, 1477-1479 (2000).

40. Hayashi,C. & and Lewis,R.V. Spider flagelliform silk: lessons in protein design, gene structure, and molecular evolution. *BioEssays* **23**, 750-756 (2001).

41. Hayashi,C. in Molecular Systematics and Evolution: Theory and Practice. eds. R.DeSalle, G.Giribet, & and W.Wheeler 209-223 (Verlag/Switzerland, Birkhaeser; 2002).

42. Hayashi,C.Y., Blackledge,T.A., & Lewis,R.V. Molecular and mechanical characterization of aciniform silk: uniformity of iterated sequence modules in a novel member of the spider silk fibroin gene family. *Mol Biol Evol* **21**, 1950-1959 (2004).

43. Hinman,M.B. & and Lewis,R.V. Isolation of a clone encoding a second dragline silk fibroin. *The journal of biological chemistry* **267**, 19320-19324 (1992).

44. Hochuli,E. in Genetic engineering, principle and methods. ed. J.K.Seltow 87-98 (Plenum Press, New York; 1990).

45. Hoogenboom,H.R., Griffiths,A.D., Johnson,K.S., Chiswell,D.J., Hudson,P., & and Winter,G. Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic Acids Research* **19**, 4133-4137 (1991).

46. Hu,X., Kohler,K., Falick,A.M., Moore,A.M.F., Jones,P.R., Sparkman,O.D., & Vierra,C. Egg Case Protein-1: A new class of silk proteins with fibroin-like properties from the spider *Latrodectus hesperus*. *J. Biol. Chem.* **280**, 21220-21230 (2005).

47. Hu,X., Vasanthavada,K., Kohler,K., McNary,S., Moore,M.F., & Vierra,C.A. Molecular mechanisms of spider silk. *Cellular and Molecular Life Sciences (CMLS)* **V63**, 1986-1999 (2006).

48. Huemmerich,D., Scheibel,Th., Vollrath,F., Cohen,S., Gat,U., & Ittah,S. Novel assembly properties of recombinant spider dragline silk proteins. *Current biology* **14**, 2070-2074 (2004).

49. Huemmerich,D., Helsen,C.W., Quedzuweit,S., Oschmann,J., Rudolph,R., & and Scheibel,T. Primary structure elements of spider dragline silks and their contribution to protein solubility. *Biochemistry* **43**, 13604-13612 (2004).

50. Inohara,N., Shimomura,S., Fukui,T., & and Futai,M. Auxin-binding protein located in the endoplasmic reticulum of maize shoots: molecular cloning and complete primary structure. *Proc. Natl. Acad. Sci. USA* **86**, 3564-3568 (1989).

51. Jin,H.-J. & and Kaplan,D.L. Mechanism of silk processing in insects and spiders. *Nature* **424**, 1057-1061 (2003).

52. Kapila,J., de Rycke,R., van Montagu,M., & Angenon,G. An *Agrobacterium*-mediated transient gene expression system for intact leves. *Plant Science* **122**, 101-108 (1997).

53. Kaplan,D., Fossey,S., Mello,M., Arcidiacono,S., Senecal,K., Multer,W., Stockwell,S., Beckwitt,R., Viney,C., & and Kerkam,K. Biosynthesis and processing of silk proteins. *MRS Bulletin* 41-47 (1992).

54. Knight,D.P. & and Vollrath,F. Biological liquid crystal elastomers. *Phil. Trans. R. Soc. Lond. B* **357**, 155-163 (2002).

55. Kovoor,J. & and Zylberberg,L. Fine structural aspects of silk secretion in a spider (*Araneus diadematus*). I. Elaboration in the pyriform glands. *Tissue and Cell* **12**, 547-556 (1972).

56. Lazaris,A., Arcidiacono,S., Huang,Y., Zhou,J.-F., Duguay,F., Chretien,N., Welsh,E.A., Soares,J.W., & Karatzas,C. Spider silk fibers spun from soluble recombinant silk produced in mammalian cells. *Science* **295**, 472-476 (2002).

57. Lewis,R.V., Hinman,M., Kothakota,S., & and Fournier,M.J. Expression and purification of a spider silk protein: a new strategy for producing repetitive proteins. *protein expression and purification* **7**, 400-406 (1996).

58. Lizardi,P.M., Mahdavi,V., Shields,D., & and Candelas,G. Discontinuous translation of silk fibroin in a reticulocyte cell-free system and in intact silk gland cells. *Proc. Natl. Acad. Sci. USA* **76**, 6211-6215 (1979).

59. Lorraine,H.L., Edmonds,D.T., & and Vollrath,F. Structural engineering of an orb-spider's web. *Nature* **373**, 146-148 (1995).

60. Lucas,F. Spiders and their silks. *Discovery* **25**, 20-26 (1964).

61. McCafferty,J., Griffiths,A.D., Winter,G., & and Chiswell,D.J. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* **348**, 552-554 (1990).

62. Mello,C.M., Senecal,K., Yeung,B., Vouros,P., & and Kaplan,D. in Silk polymers. eds. D.Kaplan, W.W.Adams, B.Farmer, & and C.Viney (American Chemical Society, Washington; 1994).

63. Menassa,R., Zhu,H., Karatzas,C.N., Lazaris,A., Richman,A., & and Brandle,J. Spider dragline silk proteins in transgenic tobacco leaves: accumulation and field production. *Plant Biotechnology Journal* **2**, 431-438 (2004).

64. Meyer,D.E. & and Chilkoti,A. Purification of recombinant proteins by fusion with thermally-responsive polypeptides. *Nature biotechnology* **17**, 1112-1116 (1999).

65. Moon,M.J. & and Kim,W.K. Ultrastructural study on the tubuliform glands in *Nephila clavata* L. Koch (Aranea: Araneidae). *Korean Arachnol.* **5** , 43-55 (1989).

66. Moon,M.J. & An,J.S. Microstructure of the silk apparatus of the comb-footed spider, *Achaearanea tepidariorum* (Araneae: Theridiidae). *Entomological Research* **36**, 56-63 (2006).

67. Motriuk-Smith,D., Smith,A., Hayashi,C., & and Lewis,R.V. Analysis of the conserved N-terminal domains in major ampullate spider silk proteins. *Biomacromolecules* **6**, 3152-3159 (2005).

68. Munro,S. & Pelham,H.R.B. A C-terminal signal prevents secretion of luminal ER proteins. *Cell* **48**, 899-907 (1987).

69. O'Brien,J.P. New structural polymers from nature. *TRIP* **1**, 228-232 (1993).

70. Olga,A., Barbara,K., Ulrike,F., Udo,C., & Klaus,D. Potato tubers as a biofactory for recombinant antibodies. *Molecular Breeding* **4**, 313-319 (1998).

71. Ortiz R., Cespedes,W., Nieves,L., Robles,I.V., Plazaola,A., File,S., & and Candelas,G. Small ampullate glands of *Nephila clavipes*. *Journal of experimental zoology* **286**, 114-119 (2000).

72. Peakall,D.B. Regulation of the synthesis of silk fibroins of spiders at the glandular level. *Comp. Biochem. Physiol.* **15**, 509-515 (1965).

73. Peakall,D.B. Regulation of protein production in the silk glands of spiders. *Comp. Biochem. Physiol.* **19**, 253-258 (1966).

74. Peters,V.H.M. Ueber der spinneapparat von *Nephila madagascariensis*. *Z. Naturforsch.* **10B**, 395-404 (1955).

75. Plazaola,A. & Candelas,G. Stimulation of fibroin synthesis elicits ultrastructural shifts in the Golgi complex of secretory cells. *PRHSJ* **9**, 227-229 (1990).

76. Plazaola,A. & and Candelas,G.C. Stimulation of fibroin synthesis elicits ultrastructural modifications in spider silk secretory cells. *Tissue and Cell* **23**, 277-284 (1991).

77. Prince,J.T., McGrath,K.P., DiGirolamo,C.M., & and Kaplan,D.L. Construction, cloning, and expression of synthetic genes encoding spider dragline silk. *Biochemistry* **34**, 10879-10885 (1995).

78. Prof.Vollrath,F. & Dr.Knight,D. in Biotechnology of Biopolymers. From synthesis to patents. eds. A.Steinbuechel & Y.Doi 873-893 (Wiley-VCH Verlag, Weinheim; 2005).

79. Reidt,W., Wohlfart,T., Ellerstroem,M., Czihal,A., Tewes,A., Ezcurra,I., Rask,L., & and Baeumlein,H. Gene regulation during late embryogenesis: the RY motif of maturation-specific gene promoters is a direct target of the FUS3 gene product. *The Plant Journal* **21**, 401-408 (2000).

80. Rising,A., Hjaelm,G., Engstroem,W., & and Johansson,J. N-terminal nonrepetitive domain common to dragline, flagelliform, and cylindriform spider silk proteins. *Biomacromolecules* **7**, 3120-3124 (2006).

81. Rosenberg,A.H., Lade,B.N., Chui,D.S., Lin,S.W., Dunn,J.J., & Sohn,S.S.H.H.a.G.S. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**, 125-135 (1987).

82. Rosenbloom,J., Abrams,W.R., & Mecham,R. Extracellular matrix 4: the elastic fiber. *FASEB J.* **7**, 1208-1218 (1993).

83. Saalbach,G., Jung,R., Kunze,G., Saalbach,I., Adler,K., & and Muentz,K. Different legumin protein domains act as vacuolar targeting signals. *The Plant Cell* **3**, 695-708 (1991).

84. Sambrook,J., Fritsch,E.F., & and Maniatis,T. Molecular cloning: a laboratory manual. (Cold Spring Harbor Laboratory Press, New York; 1989).

85. Scheibel,T. Spider silks: recombinant synthesis, assembly, spinning, and engineering of synthetic proteins. *Microbial cell factories* **3**, 14 (2004).

86. Scheller,J., Guehrs,K.-H., Grosse,F., & and Conrad,U. Production of spider silk proteins in tobacco and potato. *Nature biotechnology* **19**, 573-577 (2001).

87. Scheller,J., Henggeler,D., Viviani,A., & and Conrad,U. Purification of spider silk-elastin from transgenic plants and application for human chondrocyte proliferation. *Transgenic Research* **13**, 51-57 (2004).

88. Scheller,J. & Dr.Conrad,U. in Biotechnology of Biopolymers From Synthesis to Patents. ed. A.Steinbuechel and Y.Doi 929-943 (Wiley-VCH, Weinheim; 2005).

89. Scheller,J. & and Conrad,U. Plant-based material, protein and biodegradable plastic. *Current Opinion in Plant Biology* **8**, 188-196 (2005).

90. Sekiguchi,K. On a new spinning gland found in geometric spiders and its functions. *Annotes Zool. Jpn.* **25**, 394-399 (1952).

91. Smith,G.P. Evolution of repeated DNA sequences by unequal crossover. *Science* **191**, 528-535 (1976).

92. Sponner,A., Unger,E., Grosse,F., & Weisshart,K. Conserved C-termini of spidroins are secreted by the major ampullate glands and retained in the silk thread. *Biomacromolecules* **5**, 840-845 (2004).

93. Sponner,A., Schlott,B., Vollrath,F., Unger,E., Grosse,F., & and Weisshart,K. Characterisation of the protein components of *Nephila clavipes* dragline silk. *Biochemistry* **44**, 4727-4736 (2005).

94. Sponner,A., Unger,E., Grosse,F., & and Weisshart,K. Differential polymerisation of the tywo main protein components of dragline silk during fibre spinning. *Nature materials* **4**, 772-775 (2005).

95. Stoger,E., Ma,J.K.-C., Fischer,R., & and Christou,P. Sowing the seeds of success: pharmaceutical proteins from plants. *Current Opinion in Biotechnology* **16**, 167-173 (2005).

96. Studier,F.W. & Moffatt,B.A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113-130 (1986).

97. Szela,S., Avtges,P., Valluzzi,R., Winkler,S., Wilson,D., Kirschner,D., & Kaplan,D.L. Reduction-oxidation control of β-sheet assembly in genetically engineered Silk. *Biomacromolecules* **1**, 534-542 (2000).

98. Tillinghast,E.K. & and Christenson,T. Observations on the chemical composition of the web of *Nephila clavipes* (Araneae, Araneadea). *J. Arachnol.* **12**, 69-74 (1984).

99. Tillinghast,E.K. & and Townley,M.A. in Silk Polymers. eds. D.Kaplan, W.W.Adams, B.Farmer, & C.Viney 29-44 (American Chemical Society, Washington; 1994).

100. Tirrell,D.A. Puting a new spin on spider silk. *Science* **271**, 39-40 (1996).

101. Toepfer,R., Matzeit,V., Gronenborn,B., Schell,J., & and Steinbiss,H.-H. Expression vectors for high-level expression in dicotyledonous and monocotyledonous plants. *Meth. Enzymol.* **217**, 66-78 (1987).

102. Urry,D.W. Entropic elastic processes in protein mechanisms. I. Elastic structure due to an inverse temperature transition and elasticity due to internal chain dynamics. *Journal of protein Chemistry* **7**, 1-34 (1988).

103. Vollrath,F., Knight,D.P., & and Hu,X.W. Silk production in a spider involves acid bath treatment. *Proc. R. Soc. Lond. B* **265**, 817-820 (1998).

104. Vollrath,F. & Knight,D.P. Structure and function of the silk production pathway in the spider *Nephila edulis*. *International journal of Biological Macromolecules* **24**, 243-249 (1999).

105. Vollrath,F. & Knight,D. Liquid crystalline spinning of spider silk. *Nature* **410**, 541-548 (2001).

106. Wandelt,C.I., Khan,M.R.I., Craig,S., Schroeder,H.E., Spencer,D., & and Higgins,T.J.V. Vicilin with carboxy-terminal KDEL is retained in the endoplasmic reticulum and accumulates to high levels in the leaves of transgenic plants. *The Plant Journal* **2**, 181-192 (1992).

107. Werr,W. & and Loerz,H. Transient gene expression in a Gramineae cell line. *Mol. Gen. Genet.* **202**, 471-475 (1986).

108. Willcox,P.J., Gido,S.P., Muller,W., & and Kaplan,D.L. Evidence of a cholesteric liquid crystalline phase in natural silk spinning processes. *Macromolecules* **29**, 5106-5110 (1996).

109. Wilson,D., Valluzzi,R., & and Kaplan,D. Conformational transitions in model silk peptides. *Biophys. J.* **78**, 2690-2701 (2000).

110. Winkler,S., Szela,S., Avtges,P., Valluzzi,R., Kirschner,D.A., & Kaplan,D.L. Designing recombinant spider silk proteins to control assembley. *International journal of Biological Macromolecules* **24**, 265-270 (1999).

111. Winkler,S., Wilson,D., & and Kaplan,D.L. Controlling β-sheet assembley in genetically engineered silk by enzymatic phosphorylation/dephosphorylation. *Biochemistry* **39**, 12739-12746 (2000).

112. Xiang,C., Han,P., Lutziger,I., Wang,K., & and Oliver,D.J. A mini binary vector series for plant transformation. *Plant Molecular Biology* **40**, 711-717 (1999).

113. Xu,M. & Lewis,R.V. Structure of a protein superfiber: spider dragline silk. *Proceedings of the National Academy of Sciences* **87**, 7120-7124 (1990).

114. Yamaguchi,K., Kikuchi,Y., Takagi,T., Kikuchi,A., Oyama,F., Shimura,K., & and Mizuno,S. Primary Structure of the Silk Fibroin Light Chain Determined by cDNA Sequencing and Peptide Analysis. *J. Mol. Biol.* **210**, 127-139 (1989).

115 Zambryski,P., Joos,H., Genetello,C., Leemans,J., van Montagu,M., & and Schell,J. Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. *The EMBO Journal* **2**, 2143-2150 (1983).

116. Zhou,C.-Z., Confalonieri,F., Medina,N., Zivanovic,Y., Esnault,C., Yang,T., Jacquet,M., Janin,J., Duguet,M., Perasso,R., & and Li,Z.-G. Fine organisation of *Bombyx mori* fibroin heavy chain gene. *Nucleic Acids Research* **28**, 2413-2419 (2000).

117. Zhou,Y., Wu,S., & Conticello,V.P. Genetically directed synthesis and spectroscopic analysis of a protein polymer derived from a flagelliform silk sequence. *Biomacromolecules* **2**, 111-125 (2001).

Acknowledgements

The submitted work was performed at the Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, in group of Phytoantibody, and supported by DAAD (Deutsche Academischer Austauschdienst).

First of all, I express my thanks to PD Dr. Udo Conrad, the leader of the group, for the opportunity to work in his group, for the continuous encouragement, for the supervision of research work and the preparation of this dissertation.

I am grateful to Prof. Ulrich Wobus for allowing me to carry out this work in the Department of Molecular Genetics in the Institute of Plant Genetics and Crop Plant Research in Gatersleben.

I would like to give my thanks to Dr. Gudrun Moenke, Dr. Manfred Gahrtz, Kay Schallau, Doreen Floss, Linh Tran My, and Isabella Hermann for their willingness, their support and their helpful discussions. I am also grateful to Isolde Tillack, Ingrid Pfort, Silke Krause, Ulrike Gresch and Christine Helmold for their kind help.

My sincere thanks belong to Dr. Manfred Gahrtz and Dr. Gudrun Moenke for constructive comments on this work.

I want also to thank all present and former members of group as well as short-time visitors who all together created a positive, friendly and inspiring working atmosphere.

Finally, my gratitude belongs to my family for always supporting and respecting my decisions.

Curriculum Vitae

Name Marziya Rakhimova
Address Am Schwabeplan 1b, 06466- Gatersleben, Germany
Place of Birth Semei, Kazakhstan
Citizenship: Republic of Kazakhstan

Academic qualifications

Since July 2003	**PhD in PD Dr. Udo Conrad's group (IPK,** Gatersleben, Germany), DAAD, Leibniz-Stipendiat
1996-2003	**Research assistant** in Kazakh National Academy of Science, Institute of Molecular Biology and Biochemistry of Ajitkhozhin (IMBB), Almaty, Kazakhstan
1989-1995	**Master study in biology** (diploma thesis "Study of Natural Phenomenon of Growth of Yeast and Acidolactic-Bacteria on Whey", Biology Faculty, Kazakh State National University of Al-Farabi, Almaty, Kazakhstan

Training periods

2002	**Practice (9 month)** in phytoantibody group, IPK, Gatersleben, (Research topic: "Production of scFv against RNA dependent RNA polymerase of potato virus Y" in frame of program "Management and Environment" organised by Carl Duisberg Gessellschaft
2001	**Guest research worker (2 month)** in frame of joint-research project « Resistance to potato virus Y mediated by expression of modified RNA dependent RNA polymerase of PVY in potato plants », INTAS , Dr. Lesley Torrance, Virology unit, SCRI, UK
1998	**Training period (2 month)** in agriculture faculty, Hebrew University, Israel participation in program « Biotechnology in Agriculture : Plants and Microorganisms »

Posters and given talks

1 **M. Rakhimova** "Plant hormones long thought receptor found", given talk, *PhD student seminar, IPK, Gatersleben, April, 2006*

2 C. Münnich, **M. Rakhimova**, K. Schallau, F. Junghans, U. Spohn, U. Conrad "Spider silk proteins from plants: Production and purification", Poster, *14th international trade fair for biotechnology, Biotechnika 2005, Hannover, 20 October 2005*

3 **M. Rakhimova** and U. Conrad "Expression of spider silk elastinlike proteins in tobacco", given talk, *I st PhD student conference, IPK, Gatersleben, June, 2005*

4 **M. Rakhimova**, K. Schallau, and U. Conrad. „Expression and purification of spider silk proteins from tobacco", Poster, *Narossa Conference, Magdeburg, June, 2004*

5 Bozzhanova A. C., **Rakhimova M. G**., Karpova O. V., Akbergenov R. Zh., Polimbetova N. S., Iskakov B. K. « Analysis of expression of RNA-dependent RNA-polymerase of Potato virus Y (PVY) under regulatory elements of viruses ». Abstract book of *International Conference « Modern art of problems and achievements in genetics and sélection », Almaty, Kasakhstan, 26-27 march, 2003*

6 **Rakhimova M. G**., Bozzhanova A. S., Akbergenov R. Zh., Nazarova L. M., Nurkiyanova K. M., Iskakov B. K. « Cloning of native and modified genes of RNA-dependent RNA-polymerase of potato virus Y (PVY) under control of regulatory elements of viruses ». Abstract book of *International Symposium « MMGPB », Moscow, Russia, 18-21 november, 2001, (engl.)*

Declaration

I hereby declare that, all the work presented in this manuscript is my own, carried out with the help of the literature and aid cited.

Gatersleben, June 2007

Appendix

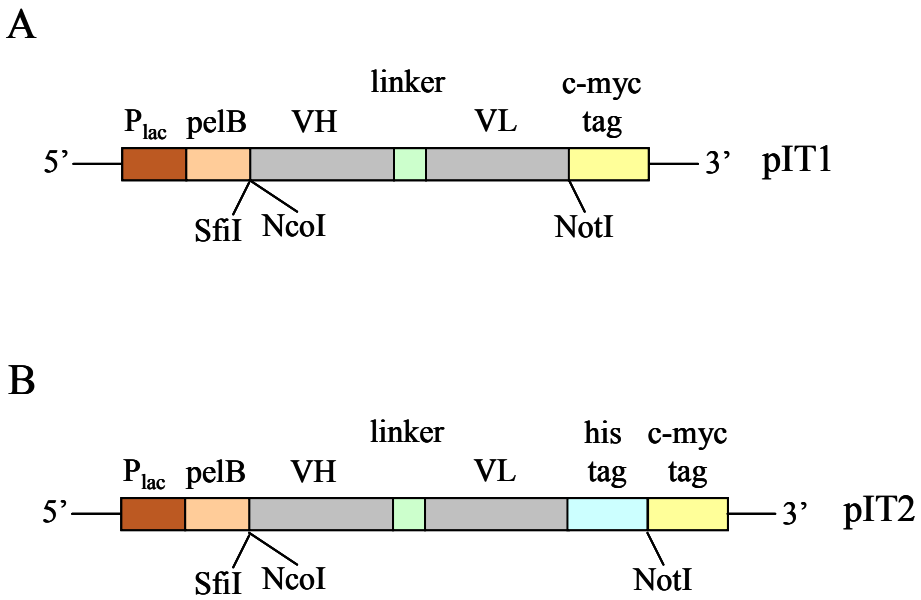


Figure 42. Schematic representation of expression cassette in phagemid vectors (A) pIT1 and (B) pIT2 harboring an scFv encoding insert. P_{lac}: Lac promoter; pelB: signal sequence; VH, linker, VL: scFv; c-myc tag: immunodetection tag; his tag: immunopurification tag; p3 CP: phage M13 p3 coat protein.

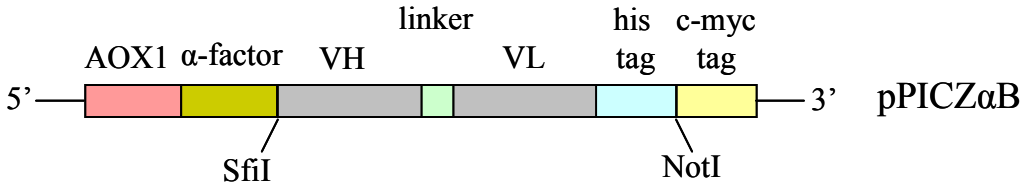


Figure 43. Schematic representation of yeast expression cassette in vector pPICZ α B harbouring an scFv encoding insert. AOX1: promoter; α-factor: signal sequence;VH,linker,VL: scFv; c-myc: immunodetection tag; his tag: immunopurification tag

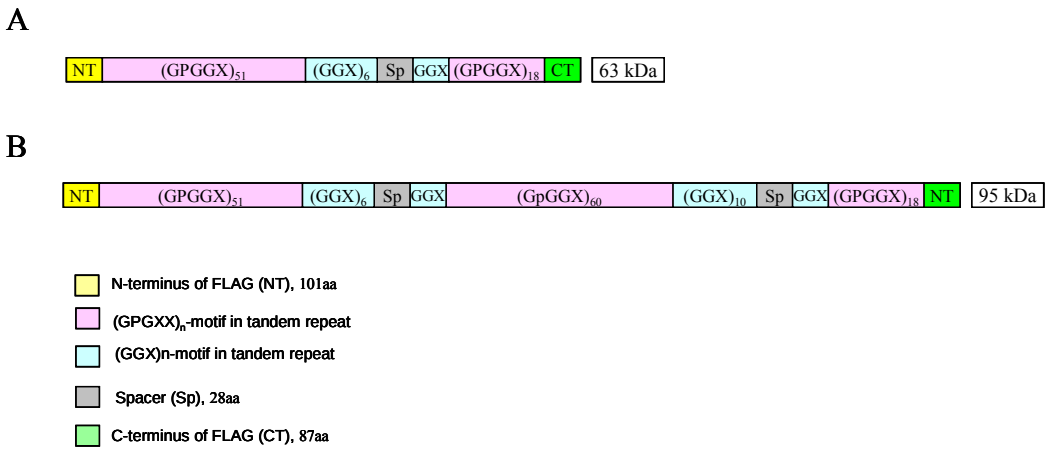


Figure 44. Schematic representation of repetitive structure of recombinant FLAG with different size: (A) one ensemble repeat, (B) two ensemble repeat

MGKGRHDTKA KAKAMQVALA SSIAELVIAE SSGGDVQRKT NVISNALRNA
LMSTTGSPNE EFVHEVQDLI QMLSQEQINE VDTSGPGQYY RSSSSGGGGG
GQGGPVVTET LTVTVGGSGG GQPSGAGPSG TGGYAPTGYA PSGSGAGGVR
PSASGPSGSG PSGGSRPSSS GPSGTRPSPN GASGSSPGGI APGGSNSGGA
GVSGATGGPA SSGSYGPGST GGTYGPSGGS EPFGPGVAGG PYSPGGAGPG
GAGGAYGPGG VGTGGAGPGG YGPGGAGPGG YGPGGAGPGG YGPGGAGPGG
YGPGGAGPGG YGPGGAGPGG YGPGGAGPGG YGPGGTGPGG YGPGGTGPGG
VGPGGAGPGG YGPGGAGPGG AGPGGAGPGG AGPGGAGPGG AGPGGYGPGG
SGPGGAGPSG AGLGGAGPGG AGLGGAGPGG AGTSGAGPGG AGPGGAGQGD
AGPGGAGRGG AGRGGVGRGG AGRGGAGRGG ARGAGGAGGA GGAGGSGGTT
*IVEDLDITID GADGPITISE ELTISGA*GGS GPGGAGPGGV GPGGSGPGGV
GPGGSGPGGV GPGGSGPGGV GPGGAGGPYG PGGSGPGGAG GAGGPGGAYG
PGGSYGPGGS GGPGGAGGPY GPGGEGPGGA GGPYGPGGAG GPYGGPGAGG
PYGPGGEGGP YGPGGSYGPG GAGGPYGPGG PYGPGGEGPG GAGGPYGPGG
VGPGGSGPGG YGPGGSGPGG YGPGGAGPGG YGPGGSGPGG YGPGGSGPGG
YGPGGSGPGG YGPGGSGPGG YGSGGAGPGG YGPGGSGPGG YGPGGSGPGG
YGPGGTGPGG TGPGGSGPGG YGPGGSGPGG SGPGGSGPGG YGPSGSGPGG
YGPSGSGPGG YGPGGSGPGG YGPGGSGAGG TGPGGAGGAG GAGGSGGAGG
SGGAGGSGGA GGSGGVGGSG GT*TITEDLDI TIDGADGPIT ISEELTISGA*
GGSGPGGAGP GGVGPGGSGP GGVGPGVSGP GGVGPGGSGP GGVGSGGSGP
GGVGPGGYGP GGSGSGGVGP GGYGPGGSGG FYGPGGSEGP YGPSGTYGSG
GGYGPGGAGG PYGPGSPGGA YGPGSPGGAY YPSSRVPDMV NGIMSAMQGS
GFNYQMFGNM LSQYSSGSGT CNPNNVNVLM DALLAALHCL SNHGSSSFAP
SPTPAAMSAY SNSVGRMFAY

Figure 45. Amino acid sequence of FLAG comprising two ensemble repeat. Underlined amino acid residues are non-repetitive parts of FLAG, red colored is putative signal peptide, cursive are spacer motifs, blue colored are (GPGGX)n pentapeptide motif, green colored are (GGX) three peptide motif.

A

```
21    L S C A A S G F T F S P Y F M P W V R Q  (40)antiFCT scFv1N2l
38    L S C A A S G F T F S N Y W M F W V R Q  (57) anti-FCT VH 3N1
39    L S C A A S G F T F S R Y A M K W V R Q  (58) anti-FCT VH 5N2
41    L S C A A S G F T F S R Y A M K W V R Q  (60) anti-FCT VH 6N2
38    L S C A A S G F T F S Y Y G M G W V R Q  (57) anti-FCT VH 7N2
36    L S C A A S G F T F S M Y P M A W V R Q  (55) anti-FCT VH 8N1
22    L S C A A S G F T F S X X X X X W V R Q  (41) Protein-CDR-Reg

Y L Q M N S L R A E D T A V Y Y C A K S Y T H F DY (66)antiFCT scFv1N2l
Y L Q M N S L R A E D T A V Y Y C A K K P Y S F DY (83) anti-FCT VH 3N1
Y L Q M N S L R A E D T A V Y Y C A K P L L P F DY (84) anti-FCT VH 5N2
Y L Q M N S L R A E D T A V Y Y C A K P L L P F DY (86) anti-FCT VH 6N2
Y L Q M N S L R A E D T A V Y Y C A K W S R H F DY (83) anti-FCT VH 7N2
Y L Q M N S L R A E D T A V Y Y C A K W N N R F DY (81) anti-FCT VH 8N1
Y L Q M N S L R A E D T A V Y Y C A K X X X X F DY (67) Protein-CDR-Reg
```

B

```
1     - - - S G F T F S A Y T M L W V R Q A P  (18) anti-FNT VH 12N1
41    C A A S G F T F S G Y G M S W V R Q A P  (60) anti-FNT VH 14N1
40    C A A S G F T F S A Y N M T W V R Q A P  (59) anti-FNT VH 15N2
41    C A A S G F T F S A Y S M H W V R Q A P  (60) anti-FNT VH 16N1
36    C A A S G F T F S R Y S M G W V R Q A P  (55) anti-FNT VH 18N2
41    C A A S G F T F S P Y I M T W V R Q A P  (60)anti-FNT scFv22N2
41    C A A S G F T F S G Y T M A W V R Q A P  (60) anti-FNT VH 23N2
24    C A A S G F T F S X X X X X W V R Q A P  (43) Protein-CDR-Reg

19     Q M N S L R A E D T A V Y Y C A K P R S S (39) anti-FNT VH 12N1
61     Q M N S L R A E D T A V Y Y C A K K R T I (81) anti-FNT VH 14N1
60     Q M N S L R A E D T A V Y Y C A K T V R A (80) anti-FNT VH 15N2
61     Q M N S L R A E D T A V Y Y C A K P R H T (81) anti FNT VH 16N1
56     Q M N S L R A E D T A V Y Y C A K M S S R (76) anti FNT VH 18N2
60     Q M N S L R A E D T A V Y Y C A K P R T T (80)anti FNT scFv22N2
61     Q M N S L R A E D T A V Y Y C A K T R K S (81) anti FNT VH 23N2
44     Q M N S L R A E D T A V Y Y C A K X X X X (64)  Protein-CDR-Reg
101    Q M N S L R A E D T A V Y Y C A K T R R I (121)antiFNT scFv21N1

40     F D Y W G Q G T L V T V S S G G G G S (58)  anti VH 12N1
82     F D Y W G Q G T L V T V S S G G G G S (100) anti VH 14N1
81     F D Y W G Q G T L V T V S S G G G G S (99)  anti VH 15N2
82     F D Y W G Q G T L V T V S S G G G G S (100) anti VH 16N1
77     F D Y W G Q G T L V T V S S G G G G S (95)  anti VH 18N2
81     F D Y W G Q G T L V T V S S G G G G S (99) anti scFv22N2
82     F D Y W G Q G T L V T V S S G G G G S (100) anti VH 23N2
65     F D Y W G Q G T L V T V S S G G G G S (83)Protein-CDR-Reg
122    F D Y W G Q G T L V T V S S G G G G S (140)anti scFv21N1

59     L A K P H T E N S F T N V W             (72)  anti VH 12N1
101    L A K P H T E N S F T N V W K D D K T L (120) anti VH 14N1
100    L A K P H T E N S F T N V W K D D K T L (120) anti VH 15N2
101    L A K P H T E N S F T N V W K D D K T L (120) anti VH 16N1
96     L A K P H T E N S F T N V W K D D K T L (115) anti VH 18N2
100    Q S I L K Y L T W Y Q Q K P G K A P K L (119) antiscFv22N2
101    L A K P H T E N S F T N V W K D D K T L (120) anti VH 23N2
84     Q S I X X X X X W Y Q Q K P G K A P K L (103)ProteinCDR-Reg
141    Q S I R T A L S W Y Q Q K P E K A P K L (160) ANTI SCFV21N1
```

Figure 46. Sequences of (A) anti FCT scFv/VH in pIT1, (B) anti FNT scFv/VH in pIT1. In red color highlighted CDR regions