Expression of Biomineralization Proteins in Bacteria and Plants

Dissertation
zur Erlangung des Grades des Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät III
Chemie, Pharmazie, Bio- und Werkstoffwissenschaften
der Universität des Saarlandes

von

Eva Weber

Angefertigt am INM - Leibniz-Institut für Neue Materialien

Programmbereich Biomineralisation

Universität des Saarlandes

2012

Tag des Kolloquiums: 19.02.2013
Dekan: Prof. Dr. V. Helms
Berichterstatter: Prof. Dr. P. Bauer, PD. Dr. habil. I. M. Weiss
Vorsitz: Prof. Dr. U. Müller
Akad. Mitarbeiter: Dr. S. Tierling

Contents

Danksagung **7**

Abstract **8**

Zusammenfassung **9**

Summary **10**

Abbreviations **12**

1 Introduction **14**

1.1 Biomineralization: Lessons from Nature 14

1.1.1 Motivation - Why Biomineralization? 14

1.1.2 Model Organisms for Biomineralization 14

1.1.3 The Impact of Organic Macromolecules on Mineral Formation . . 17

1.1.4 Biomineralization Proteins 18

1.2 Concepts of Mineralization 20

1.3 Calcium Biomineralization of Land Plants 21

1.4 The Structure of the Plant Cell Wall 22

1.4.1 Cellulose 23

1.4.2 Hemicelluloses 23

1.4.3 Pectins 24

1.4.4 Cell Wall Proteins 25

1.5 Biotechnology - Plants as Bioreactors for Protein Biosynthesis 26

1.5.1 The Advantages of Plants as Bioreactors 26

1.5.2 Strategies and Examples for Xenogenic Protein Biosynthesis in Plants 28

1.5.3 Plant Secretory Pathway 30

1.5.4 Misfolded Protein Degradation System 31

1.6 Objectives of the work 32

2 Material and Methods **34**

2.1 Materials 34
2.1.1 Plant Materials 34
2.1.2 Bacterial Strands 34
2.1.3 Plasmids 35
2.1.4 Kits 40
2.1.5 Chemicals 40
2.1.6 Antibiotics 42
2.1.7 Enzymes 42
2.1.8 Antibodies 43
2.1.9 Diverse Materials 43
2.1.10 Oligonucleotids 44
2.1.11 Gen Synthesis and Sequences 44
2.1.12 Sequencing 45
2.1.13 Solutions and Buffers 45
2.1.14 Molecular Weight Standards 50
2.1.15 Equipment 50
2.1.16 Software and Web Tools 51
2.2 Methods 52
2.2.1 Molecular Biological Methods 52
2.2.2 Microbiological Methods 60
2.2.3 Protein Methods 62
2.2.4 Plant Biological Methods 70
2.2.5 Spectroscopic Methods 76
2.2.6 Microscopic Methods 77
2.2.7 Calcium Carbonate Precipitation Assays 78

3 Results **80**

3.1 Expression of Biomineralization Proteins in Bacteria 80
3.1.1 Cloning Strategy 80
3.1.2 Protein Expression and Purification 82
3.1.3 Protein Characterization 85
3.1.4 Activity Assay of Recombinant Proteins 90
3.1.5 Synthesis of GFP-Silica Nanoparticles 104

3.2 Expression of Biomineralization Proteins in Plants 105
3.2.1 Selection of the Signal Peptides for Cell Wall Delivery 105
3.2.2 Cloning Strategy - Change of Resistance Gene 106
3.2.3 Expression of Perlucin-GFP in *Arabidopsis thaliana* 109
3.2.4 Summary: Expression of Perlucin-GFP in *Arabidopsis thaliana* . . 117
3.2.5 Expression of Biomineralization Proteins in *Nicotiana benthamiana* 119
3.2.6 Summary: Expression of Biomineralization Proteins in *Nicotiana benthamiana* 123

4 Discussion and Outlook **124**
4.1 Expression and Purification of Biomineralization Proteins in Bacteria . . 124
4.1.1 Characterization of Recombinant GFP, GFP-Perlucin and Perlucin 124
4.1.2 Influence of Recombinant GFP and GFP-Perlucin on Calcium Carbonate Formation 126
4.1.3 Characterization of Metastable GFP and GFP-Perlucin Vaterite Composites 128
4.1.4 Summary and Outlook 129
4.2 Expression of Biomineralization Proteins in Plants 130
4.2.1 Expression of Perlucin-GFP in *Arabidopsis thaliana* 131
4.2.2 Possible Effects of Perlucin-GFP Expression on Plant Development 132
4.2.3 Expression of Perlucin-GFP, N16N-GFP and OC17-GFP in *Nicotiana benthamiana* 136
4.2.4 Summary and Outlook 138

Bibliography **154**

5 Appendix **155**
5.1 DNA Sequences 155
5.2 Expression of Biomineralization Proteins in Bacteria 158
5.2.1 Purification Steps of Recombinant Proteins 158
5.2.2 SDS-PAGE Analysis of Recombinant Proteins 159
5.2.3 Western Blot Analysis of GFP-Perlucin 160
5.2.4 Microscopical Analysis of Calcium Carbonate Precipitates 161
5.2.5 Modelling of Calcium-GFP Interaction 161
5.2.6 Comparison of Batch and Chromatography Purification 163
5.2.7 Field-Flow Fractionation - Protein Agglomeration Analysis 165

5.3 Expression of Perlucin-GFP in *Arabidopsis thaliana* 167
5.3.1 Genotyping of the T2 Perlucin-GFP Plants 167
5.3.2 RNA Extraction of T2 Plants 168
5.3.3 Germination Rate and Root Growth Assay of T2 Generation . . . 169
5.3.4 Genotyping T3 Plants 170
5.3.5 Seed Analysis of T3 Plants 172
5.4 Expression of Biomineralization Proteins in *Nicotiana benthamiana* . . . 173

Publications **174**

Danksagung

Im Besonderen möchte ich mich bei Herrn Prof. Eduard Arzt, PD Dr. habil. Ingrid Weiss und Frau Prof. Petra Bauer bedanken, die mir die Möglichkeit gegeben haben, diese Arbeit im Arbeitskreis Biomineralisation unter der Leitung von Frau PD Dr. habil. Ingrid Weiss am INM anzufertigen. Hierbei danke ich Prof. Petra Bauer für die Übernahme der Erstkorrektur und Herrn Prof. Eduard Arzt für die Übernahme der Zweitkorrektur. Die Arbeit wurde von PD Dr. habil. Ingrid Weiss betreut und ich danke Ihr im Besonderen für die gute Betreuung, die vielen hilfreichen Diskussionen und Ihre Unterstützung während der gesamten Zeit.
Magdalena Eder, Christina Guth, Sarah Schmitt, Angela Rutz und Sabrina Schleich möchte ich für Ihre Freundschaft, kontinuierliche Unterstützung und Hilfsbereitschaft danken. Dabei möchte ich besonders Magdalena Eder und Christina Guth für die schöne gemeinsame Zeit und Ihre Geduld danken.
Gabi Klein und Dr. Tobias Kraus möchte für Ihre Unterstützung bei der Etablierung der Feld-Fluss Fraktionierung danken.
Für die gute Unterstützung beim Planen und Einrichten des Pflanzenanzuchtsbereiches möchte ich der gesamten E-Werkstatt und der Haustechnik danken.
Des Weiteren möchte ich mich bei allen ehemaligen und derzeitigen Kollegen vom INM herzlich bedanken, die meine Zeit am INM einzigartig gemacht haben: Marleen Kamperman, Ramgopal Balijepalli, Dadhichi Paretkar, Florian Hausen, Ralf Muth, Karsten Moh, Felix Wählisch und Christian Cavelius, Katrin Brörmann, Maurizio Micciche, Christian Held, Regina Gratz, Anna Kreuder und Gabriele Koster.
Für das Korrekturlesen meiner Arbeit geht mein Dank insbesondere an Magdalena Eder, Archana Balijepalli, Sabrina Schleich, Phil Egberts, Julia Callender und Martina Bonnard.
Für die gute Unterstützung in pflanzen-biologischen Fragen möchte ich meiner früheren Arbeitsgruppe und hier insbesondere Angelika Anna und Johannes Meiser danken. Einen ganz besonderen Dank gilt meinen Eltern, die mich immer unterstützt und diese Arbeit erst ermöglicht haben.

Abstract

The fascinating properties of natural biominerals, such as stiffness combined with toughness, are difficult to mimick in synthetic pathways. Therefore, this work investigated the biosynthesis of biomineralization proteins in different organisms for producing synthetic minerals under the control of a protein phase. The soluble nacre protein perlucin from *Haliotis* was produced in bacteria but remained insoluble after native extraction. It was partially solubilized when fused to a GFP, which was additionally identified to be a strong inhibitor of calcium carbonate formation. Various forms of synthetic biominerals were obtained in simple calcium carbonate precipitation assays depending on the first ionic interaction partner, assay volume and protein concentrations. In comparison to bacteria as less complex organisms, the biomineralization proteins perlucin, N16N and ovocleidin-17 were also expressed in the plant cell wall of *Arabidopsis* and *Nicotiana*, which has not been described in the literature so far. This work demonstrated that N16N-GFP and ovocleidin-17-GFP can be expressed in epidermal leaf cells of *Nicotiana*. In perlucin-GFP transformed *Arabidopsis*, the plant morphology changed although neither the protein nor RNA were detected. The expression of foreign biomineralization proteins in biotechnologically relevant plants was demonstrated for the first time, opening routes toward new and improved materials.

Zusammenfassung

Biomineralien mit ihren herausragenden Eigenschaften sind schwierig in synthetischen Systemen herzustellen. Daher wurde in dieser Arbeit die Herstellung von Biomineralisationsproteinen in verschiedenen Organismen untersucht, mit dem Ziel künstliche Biomineralien unter der Kontrolle einer Proteinphase zu synthetisieren. Das lösliche Perlmutt-Protein Perlucin aus *Haliotis* wurde im bakteriellen System hergestellt. Obwohl Perlucin aus Bakterien unlöslich war, konnte es durch Fusion an GFP nativ extrahiert werden. Dabei wurde entdeckt, dass GFP inhibierend auf die Bildung von Calciumcarbonat wirkt. Hiermit konnten verschiedene synthetische Biomineralien in Abhängigkeit des ersten ionischen Interaktionspartners, Reaktionsvolumen und Proteinkonzentration erzeugt werden. Vergleichend zu dem bakteriellen Expressionssystem wurden die Proteine Perlucin, N16N und Ovocleidin-17 in der pflanzlichen Zellwand von *Arabidopsis* und *Nicotiana* exprimiert. Dieser Ansatz ist derzeit in der Literatur noch nicht beschrieben. Es wurde gezeigt, dass N16N-GFP und Ovocleidin-17-GFP in *Nicotiana* exprimiert werden. Obwohl von Perlucin-GFP weder Protein noch RNA nachgewiesen werden konnten, traten in *Arbabidopsis* morphologische Veränderungen auf. Die erfolgreiche Expression von Biomineralisationsproteinen in biotechnologisch relevanten Organismen bietet eine Grundlage für die Herstellung von neuen und verbesserten Biokomposit-Materialien.

Summary

Natural materials such as mollusc shells, wood and plant cell walls have fascinated scientists due to their extraordinary properties which were optimized with respect to different functions during million of years of evolution. Despite their evolutionary different origin, mollusc shells and plant cell walls have similarities in certain aspects such as the hydrogel texture and the composition of biopolymers. Many functionalities are the result of hierarchical textures, consisting of different organic and inorganic phases. These specific interactions on different length scales have major impact on material properties. How organisms control organic and mineral formation for the synthesis of hierarchical composites is not fully understood so far. Despite many well known proteins involved in biomineralization, only a few are characterized to some extent such as the protein (used here) encoding sequences for N16N (mollusc shell, *Pinctada fucata*), Ovocleidin-17 (eggshell, *Gallus gallus*) and perlucin (mollusc shell, *Haliotis laevigata*).

To combine evolutionary different system, this work dealt with the biosynthesis of biomineralization proteins in pro- and eukaryotic systems and aimed to synthesize minerals under the control of protein phases. Biomineralization proteins were selected with respect to their ability to induce mineral formation. All proteins were combined with a green fluorescent protein (GFP) for protein localization during interaction with the mineral phases *in vitro* or directly in organisms. Two strategies were pursued in this work.

In the first approach, the native soluble protein perlucin was produced in bacteria. Bacteria are relatively simple organisms and often used for the biotechnological synthesis of proteins. The recombinant proteins were used to study mineral-protein interaction *in vitro*. In the second part, the proteins perlucin, N16N and OC17 were introduced into plant organisms. The cell wall was used as the target compartment because of its important functions in mechanically supporting the plant cell and mimicking the natural biomineralization matrix. This approach offers the opportunity to modify the texture of plant tissues by the organism itself. It has the advantage that the hierarchies and nanoscale modifications are directly provided and preserved by the organisms which

cannot be achieved in such a complexity by means of chemical approaches.

In the bacterial expression system, a high protein yield was obtained after optimization of the growth conditions and protein purification processes. It was shown that GFP was soluble under native conditions whereas the recombinant perlucin was found to be insoluble. This insolubility was partly overcome through the fusion of perlucin to the GFP domain. Further, the influence of the recombinant protein on calcium carbonate precipitation was investigated in two independent assays which enabled the synthesis of biominerals *in vitro*. Surprisingly, GFP was identified to be a strong inhibitor of calcium carbonate precipitation under certain conditions. At very low protein concentrations, the protein was able to interact with metastable calcium carbonate phases. Various forms of synthetic biominerals were obtained in dependence on the first ionic interaction partner, assay volume and protein concentrations. Among these composites, crystals were formed with a lamellar mineral/protein arrangement similar to the texture of nacre.

In the plant biotechnological system, two model species were selected for a comparative analysis. Tobacco was selected as system for the temporary expression of perlucin-GFP, N16N-GFP and OC17-GFP in epidermal leaf cells. *Arabidopsis thaliana* was used for a permanent expression of perlucin-GFP in the whole plant organism in order to investigate all developmental stages of the plant. This work demonstrated for the first time that the expression of N16N-GFP and OC17-GFP was achieved in tobacco epidermal leaf cells. Perlucin-GFP was so far not detected on the protein and mRNA level, neither in tobacco nor in *Arabidopsis*. However, effects such as reduced germination rate and differences in root length growth were observed in perlucin-GFP *Arabidopsis thaliana* lines. These results indicate an interference of perlucin with plant cellular processes.

This study revealed the complexity of the different expression systems regarding the synthesis of biomineralization proteins. The possibility to use biotechnologically produced proteins for material synthesis may have a high impact on the production of biocomposite materials in the future. The functional integration of biomineralization proteins in the hierarchal texture of cell walls opens a versatile route to modify the mechanical properties of specialized plant organs. This could lead to biopolymers-composites, which are more accessible for downstream processing and would be of high importance for industrial approaches such as paper industry. Therefore, this work opens biotechnological and bioinspired routes towards new and improved materials.

Abbreviations

ACC	Amorphous Calcium Carbonate
AF4	Asymmetric Field-Flow-Fractionation System
AGP	Arabinogalactan Proteins
AP	Alkaline Phosphatase
bp	Base Pair
BP	Biomineralization Protein
BSA	Bovine Serum Albumin
CBP	Chitin Binding Protein
CLSM	Confocal Laser Scanning Microscopy
ConA	Concanavalin A
CV	Column Volume
DIC	Differential Interference Contrast
DLS	Dynamic Light Scattering
DNA	Desoxyribonucleic Acid
DTT	Dithiothreitol
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediaminetetraacetic Acid
ERAD	ER Associated Degradation System
FFF	Field-Flow-Fractionation
FITC	Fluorescein Isothiocyanate
GFP	Green Fluorescent Protein
GRP	Glycin Rich Protein
Gm	Gentamycin Resistance
H_2O_{MP}	H_2O Millipore
HA	Haemagglutinin

HG	Homogalacturonan
HPLC	High Pressure Liquid Chromatography
HPRP	Hydroxy Proline Rich Protein
HIS_6	Histidin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Km	Kanamycin Resistance
LB	Left Border
LC	Liquid Crystal
LS	Light Scattering
MCS	Multiple Cloning Site
MS	Murashige and Skoog
O.D.	Optical Density
PCR	Polymerase Chain Reaction
PER	Ammonium Persulphate
PME	Pectin Methylesterase
PTM	Posttranslational Modification
PRP	Prolin Rich Protein
QELS	Quasi Elastic Light Scattering
RB	Right Border
RER	Rough Edoplasmic Reticulum
RGI	Rhamnogalacturonan I
RGII	Rhamnogalacturonan II
Rh	Hydrodynamic Radius
RT	Room Temperature
SDS-PAGE	SDS Polyacrylamide Gel Electrophoresis
SP	Signal Peptide
SRP	Signal Recognition Peptide
TEM	Transmission Electron Microscopy
TGN	Trans Golgi Network
UPR	Unfolded Protein Response
V-ATPase	Vacuolar H^+-ATPase
V-PPase	Vacuolar H^+-Pyrophosphatase

1 Introduction

1.1 Biomineralization: Lessons from Nature

1.1.1 Motivation - Why Biomineralization?

Biomineralization describes the process of mineral formation by organisms. The formed biominerals have fascinated scientists due to their extraordinary properties such as light weight combined with mechanical strength. It is not fully understood and was not achieved in abiotic systems so far, how organisms are able to control mineral shape and polymorph in a remarkable arrangement over different length scales [Lowenstam1989]. One of the major challenges is to follow the growth of bio-composites and to apply these fundamental concepts on the synthesis of new and improved materials. This work investigates the formation of artificial bio-composites due to the interaction of biotechnologically produced biomineralization proteins with calcium carbonate and reports for the first time the expression of non-plant biomineralization proteins in plant organisms, opening routes toward new and sustainable materials. In the absence of a clear definition for the term "biomineralization proteins", it has to be mentioned that the term is used in this work for proteins, which are related to biomineralization processes. In addition, the term "artificial or synthetic biomineral" is used for the product of mineral interaction with biomineralization proteins *in vitro*.

1.1.2 Model Organisms for Biomineralization

Biomineralization is today a widespread phenomenon among different phyla and various forms of biominerals exists [Lowenstam1989]. The first macroscopic biomineralization events can be traced back to 570 million years ago and had a major impact on living systems with the ability to form hard skeletal structures [Mann2001]. Biominerals fulfill a variety of functions and a very common one is gravity sensing in land and sea animals (such as otoliths). Bacteria use single crystals of magnetite (Fe_3O_4) for navigation by means of the earth's magnetic field whereby remarkable structures were formed

by diatoms that produce skeletons of amorphous silica [Mann2001]. Calcium within calcium carbonate is a very widespread ion for biomineral formation due to its high stability within a biological environment [Mann2001]. Model organisms for calcium carbonate mineralization have been the main focus of this thesis, including mollusc shell and eggshell, and their morphology is described in the following section.

1.1.2.1 Mollusc Shell-Structure and Properties

One well studied model organism for calcium carbonate biomineralization is the mollusc shell consisting of an outer calcite and an inner aragonite layer (nacre) [Addadi2006]. Aragonite and calcite are polymorph forms of calcium carbonate. Nacre is formed from single crystals of aragonite (95 % of nacre weight) in a highly oriented form and they are separated by thin layers of organic matrix (5 % of nacre weight) as shown in Figure 1.1.

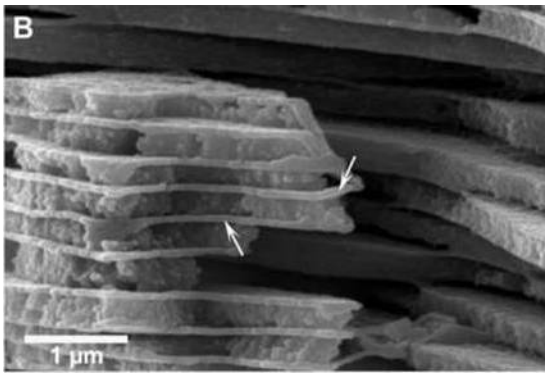


Figure 1.1: Scanning electron microscopy (SEM) image from the cephalopod *N. pompilius* shell. Thin 30 nm layers of matrix (arrows) separate the 500 nm aragonite platelets in parallel to the shell surface. The platelets are arranged in vertical rows reaching a final size of 10 μm x 10 μm x 0.5 μm [Addadi2006].

The enhanced toughness of the mollusc shell compared to pure inorganic calcium carbonate was first analyzed by mechanical approaches in 1974 [Currey1974]. Its reduced sensitivity to internal defects caused by external forces is today explained by the composite structure [Evans2001]. The additional hierarchical architecture is also a key feature that helps to explain the extraordinary properties [Laraia1989]. Furthermore, the scale of the parameters has to be considered, both in the natural system [Mayer2005] and for biomimetic artificial nacre materials [Tang2003], [Deville2006]. Amorphous calcium carbonate (ACC), which is unstable in its pure form, was shown to be stabilized or used

as transient precursor phase for its transition to stable calcite or aragonite minerals in organisms [Addadi2003].

1.1.2.2 Eggshell

The growing mollusc shell represents a slow mineralization process compared to the fast formation of eggshells in a time scale of ∼24 h. The egg from *Gallus gallus* is assembled in specific regions of the hen oviduct. First, the compounds of the egg white are deposited (2-3 h), followed by the inner and outer shell membranes (1-2 h), which consist of a fibrous network of collagen [Wong1984], [Arias1991]. Subsequently, organic aggregates are deposited on the outer membrane, serving as nucleation sites for calcification (16-17 h) [Nys2004], [Rose2009].

A SEM image of a cross-fractured eggshell revealed the shell membrane (Figure 1.2, SM), the palisade layer (PL) and the mammillary layer (ML). The ML layer (approximately 100 μm thick) contains spherulitic calcite and is attached in knob-shaped end-structures to the collagen fibres of the outer membrane whereby the inner membrane remains non-calcified [Arias1993], [Nys2004], [Rose2009].

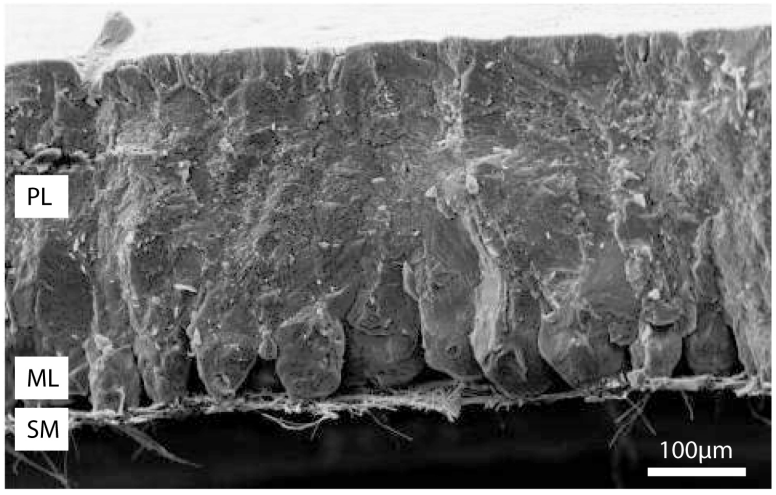


Figure 1.2: SEM image of a cross-fractured eggshell with the shell membrane (SM), mammillary layer (ML), palisade layer (PL). The ML consists of spherulitic calcite crystals and the PL contains elongated calcite crystals, which are attached via knob-shaped end-structures to the collagen fibrous network of the outer SM [Rose2009]. Image has been modified from [Dobiasova2003].

1.1.3 The Impact of Organic Macromolecules on Mineral Formation

Today, it is widely accepted that the influence of the organic matrix is essential for the control of mineral formation, although the specific composition and its interactions with the mineral phase is not fully understood. Organisms synthesize biominerals either in a "biologically controlled" or "biologically induced" way, which represents the two main strategies for biomineralization [Lowenstam1981], [Mann1983]. The latter refers to the interaction of aqueous environment with biological activity, such as the release of particular ions by the cells, with less control on mineral formation [Lowenstam1981], [Mann2001] and is not further discussed here.

The "biologically controlled" mineralization refers to processes of cellular activities such as nucleation under direct control of the organism by the influence of macromolecules and the formation of a saturated solution in a separated space as precondition for mineralization [Lowenstam1981], [Mann1983]. The macromolecules can be categorized in a water-soluble and a water-insoluble fraction, whereby the insoluble fraction acts as framework for controlled mineralization [Falini1996].

In vitro assays revealed the ability to control nucleation of the calcium carbonate polymorph aragonite by means of pre-adsorption of macromolecules on a β-chitin-silk substrate [Falini1996]. The function of the insoluble matrix as a two- or three-dimensional template was shown for the growth of aragonite crystals, demonstrating the impact of the insoluble fraction on crystal formation [Heinemann2006]. In former studies, the influence of single soluble molecule species such as phosphorylated amino acids, acidic sulphated polysaccharides and acidic amino acids on mineral formation was shown *in vitro* [Lee1977], [Addadi1985], [Lowenstam1989]. The interaction of these proteins with calcium seems to play a major role in a cross-linking process to hydrophobic macromolecules such as β-chitin in mollusc shells and facilitates the formation of β-sheet protein structures, serving as template for oriented nucleation [Worms1986]. A model of a demineralized nacreous layer is shown in Figure 1.3, consisting of β-chitin, aspartic acid rich glycoproteins and a silk fibroin gel matrix to pre-fill the site of mineralization [Levi-Kalisman2001].

Furthermore, the control over polymorph selection by the organic matrix [Belcher1996], [Falini1996] with the transition of precursor phases was shown in several cases [Gotliv2003]. Some organisms are able to produce stable vaterite [Lowenstam1989], for example the freshwater bivalve *Corbicula fluminea* [Frenzel2011], [Spann2010], whereas the crayfish uses small metabolites for stabilizing ACC [Sato2011]. ACC and vaterite likely transform into aragonite and calcite under ambient conditions [Beck2010], [Lippmann1973].

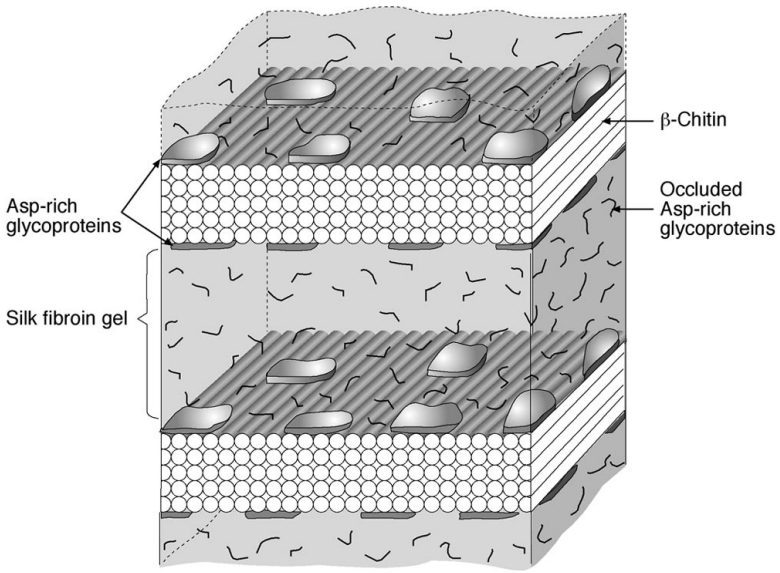


Figure 1.3: Schematic representation of the organic matrix of a demineralized mollusc shell from *Atrina*. Chitin builds the framework with patches of acidic macromolecules, which may serve as nucleation sites for mineral formation and the silk fibroin gel matrix acts as space-holder for the growing mineral phase [Levi-Kalisman2001].

Today, the influence of the microenvironment [Levi1998] and the organic matrix with tunable solubility is widely accepted as the major driving force for controlled biomineralization in a wide range of phyla [Lowenstam1981], [Mann2001], [Weiner2009]. However, the detailed composition of the environment with defined molecule interaction is hard to investigate *in vivo*, but *in vitro* studies with extracted organic species indicate certain aspects regarding their role and activity within biomineralization processes.

1.1.4 Biomineralization Proteins

Various biomineralization proteins are purified from the organic matrix of different organisms or their sequences deduced from cDNA/expression libraries. They were analyzed regarding their modular structure and multifunctionality such as enzymatic or structural activity [Weiss2008]. A well studied protein is nacrein having a carbonic anhydrase (enzymatic) and calcium binding (structural) function as demonstrated *in vitro* [Miyamoto1996]. For fundamental research and biomimetic approaches, these proteins offer a rich source of inspiration. The proteins perlucin, ovocleidin-17 and N16N, which were in the main focus of this work, will be discussed in more detail in the next sections.

1.1.4.1 Perlucin from Haliotis laevigata

Perlucin was the first soluble protein isolated from the nacre of the mollusc *H. laevigata* [Weiss2000]. It contains 155 amino acids and has a molecular weight of 17 kDa and corresponds to the N-terminal part of a precursor protein with 240 amino acids [Dodenhof2010]. The native perlucin protein is glycosylated (asparagine residue) [Weiss2000],[Mann2000] and incorporates into calcium carbonate *in vitro* [Blank2003]. Perlucin contains a C-type lectin domain (calcium dependent) with six cysteines, which may form disulfide bonds. It has a broad sugar-binding range including mannose and galactose residues, followed by two C-terminal repeat sequences, each with 10 amino acids [Weiss2000], [Mann2000]. The functionality of perlucin as nucleation promotor of calcium carbonate [Weiss2000] and the interaction of perlucin with calcite crystals inducing the growth of new mineral layers was demonstrated *in vitro* [Blank2003].

Recombinant forms of perlucin were produced using prokaryotic expression systems, however, purification has only been possible so far when including a denaturing polyacrylamide gel electrophoresis step [Blohm2007]. Subsequent gentle dithiothreitol (DTT) renaturing steps recovered the mineralization activity of perlucin with respect to calcite nucleation [Blohm2007]. Wang and colleagues achieved natively purified forms of perlucin (*Haliotis discus discus*) by recombinant fusion to maltose binding protein and demonstrated their interference with calcite crystals [Wang2008]. Former studies showed that the N-terminus of perlucin is a promising candidate to overcome inhibitory effects on mineralization, which were reported for natively extracted nacre biomineralization protein mixtures and which occured therein in a protein concentration dependent manner [Wheeler1981],[Heinemann2011].

1.1.4.2 N16N from Pinctada fucata

N16N corresponds to the first 30 amino acids of the N-terminal part of N16-1 (here called N16) [Samata1999], [Metzler2010]. N16 is one of three proteins, which have been extracted from the water insoluble matrix of the nacreous layer of the bivalve *Pinctada fucata* [Metzler2010]. The pif proteins (pif97, pif80), which are key regulator proteins for aragonite formation in *Pinctada fucata* were suggested to form a multi-protein complex with N16 for directed synthesis of aragonite layers [Suzuki2009], [Metzler2010]. N16N is rich in aspartate and glutamate residues (5 (16.6 %) negatively charged amino acids) and was suggested to bind calcium [Collino2008], [Metzler2010]. Interaction of Ca^{2+} ions with N16N changed the protein chain dynamics as well as the local 3-dimensional

protein structure and allowed the protein to adsorb on a calcite surface [Collino2008]. A self-assembly of synthetic N16N into layers promoted the growth of aragonitic crystals [Metzler2010] and after pre-adsorption on β-chitin, aragonite and vaterite or ACC was formed *in vitro* [Ponce2011]. In a similar crystallization experiment, combining a silk fibroin hydrogel and N16N on β-chitin, resulted in a metastable vaterite and ACC deposition with hemispherical centers [Keene2010]. The native N16-3 (isoform 3) exhibited a pH-dependent oligomerization and a tendency to an ordered protein secondary structure at high calcium concentrations [Ponce2011]. In particular, the N-terminal sequence of N16 has the potential to influence mineralization process *in vivo* and *in vitro* [Metzler2010].

1.1.4.3 Ovocleidin-17 (OC17) from Gallus gallus, Eggshell

Compared to the relatively slow calcification of mollusc shell, the eggshell is formed within a time period of approximately 24 hours. The biomineralization protein OC17 was first purified by Hincke and coworkers (1995) who were also the first to report on its potential for eggshell formation [Hincke1995]. The protein is expressed in the tissue from where the mammillary and palisade layers originate (Figure 1.2). Furthermore, morphological changes of calcium carbonate precipitates in the presence of OC17 were observed [Reyes-Grajeda2004], [Marin-Garcia2008]. Compared to perlucin, the sequence of OC17 revealed similarities to a C-type lectin-like domain, however, the lectin-binding capacity has not been demonstrated so far and recently, simulation-based analysis demonstrated that OC17 is less likely involved in controlling crystal morphology but may promote calcite nucleation [Freeman2010], [Freeman2012].

1.2 Concepts of Mineralization

To understand how biomineralization processes are controlled at the cellular level, basic concepts, such as nucleation and growth of mineral particles and crystals, have to be considered. One of these key aspects is the solubility of an inorganic salt defined as the number of moles of a pure solid which is dissolved in a liter of solvent at a specific temperature. The physical parameters for mineral solubility have to be controlled by the organism to prevent unwanted precipitation of the mineral phase. The solvent-solute interactions have to take into account both the energy and the entropy [Mann2001]. Supersaturation is another key aspect in biomineralization and describes

the extent of a solution to be out of the equilibrium. The supersaturation is a major driving force for inorganic precipitation. The organic matrix can induce heterologous nucleation in a supersaturated solution due to the formation of crystal nuclei on their surface. The supersaturation can be controlled by organisms via ion gradients in such a manner that the activation energy for nucleation is reduced with the increase in supersaturation [Mann2001]. Different energy barriers have to be overcome during crystal growth, which influences the kinetic of the reaction. The role of additives influencing the crystal growth is manifold to such an extent that they adsorb on specific crystal faces. The crystal morphology can be strongly changed due to the molecular-specific interaction of mineral surfaces with the additives. Therefore, the interference of additives with the mineral surface can also inhibit or terminate crystal growth. For more details of concepts in mineralization see [Mann2001]. To what extent and how exactly these mineralization concepts are implemented by organisms by means of biomolecules and membrane compartments is currently under debate. The formation of biominerals is not only known for animals but also for plant organisms, which is described in the next section.

1.3 Calcium Biomineralization of Land Plants

Biominerals are widely distributed in higher plants with calcium oxalate as the most prominent form. The formation of crystals as osmotic inactive salt serves as calcium storage in a non toxic form, but can also act as physical defense against herbivores [Arnott1970]. Calcium oxalate is deposited in various cell compartments such as the cell wall [Hudgins2003] or specialized cells, also called idioblasts [Foster1956], [Franceschi2005]. Much descriptive information is available on the crystal morphologies, but little is known about the cellular processes involved in crystal formation. Fundamental contribution offered the work of Nakata and co-worker, showing a modification in crystal shape and size due to a point mutation in the *cod* gene (calcium oxalate defective) in *Medicago truncatula* [Nakata2000],[Franceschi2005]. Different mutants were identified, all with modified calcium oxalate crystals implicating a genetic regulation by the plant organisms on calcium regulation. Bouropoulos et al. (2001) extracted biomineral associated macromolecules from tobacco and showed their promoting influence on calcium oxalate nucleation in a concentration of 4 μg/ml *in vitro* by means of a turbidity measurement of the solution [Bouropoulos2001].

Furthermore, plants contain calcium carbonate biominerals in the form of ACC, which

is rather common compared to single crystal formation [Arnott1970]. Calcium biominerals are often embedded in specialized cells, so-called cystoliths, which are integrated in the cell wall in leaves of different plant species such as *Ficus* and *Helianthus* plants [Arnott1970]. The cystolith formation was investigated in the mulberry tree in various studies [Sugimura1998],[Sugimura1999], [Nitta2006], [Sugimura2007], [Katayama2008]. The amorphous and very soluble character of the calcium carbonate in cystoliths is considered to have a physiological importance in calcium homeostasis of the plant [Nitta2006].

Crystalline calcium carbonate and calcium oxalate were observed in the model plant organism *Nicotiana* (tobacco). Various types of calcium associated biominerals (oxalate, carbonate (aragonite and calcite)) were identified under metal ion stress after extrusion via trichomes [Sarret2006]. It was shown that zinc interacts with the biominerals and that the zinc stress tolerance of the plants increases with the amount of supplied calcium. The detoxification under heavy metal stress in *Nicotiana* was also analyzed by several groups ([Harada2008], [Choi2005], [Isaure2010]). The formation of biominerals is well described in plants, but little is known about the genetic regulation of these processes.

1.4 The Structure of the Plant Cell Wall

Most biominerals represent a specialized kind of extracellular matrix, which protects organisms from unfavourable environmental conditions and gives them their skeletal tissue shape. Similarily, plant cells are surrounded by a specialized extracellular matrix, which has comparable functions in support and protection. The structure and composition of the cell wall is described in the following sections. The cell wall is one of the main characteristic elements, which distinguish animals from plants and take over a diversity of functions such as cell shape [Bacic1988], defense against pathogens (summary [Brett1996]), carbon-source [Franco1996], mechanical strength and flexibility, in particular during cell expansion. The growing cell wall has a typical pH between 4.5 and 6 and is subjected to the expansin activity, which is induced via acidification. This low pH is under the control of plasma membrane proton pumps (H^+-ATPase) and induces cell wall alteration for facilitated cell expansion [Cosgrove2005].

The cell wall framework consists of cellulose, the most abundant biopolymer on earth, which serves as important raw material for industry, such as textile or paper production. One major bottleneck presents the down-stream processing of cells for industrial applications in an eco-friendly and economic way, because some cell wall components

bear a high complexity, stability and resistance against chemical agents. There are two main types of cell walls in plants: The primary cell wall and the secondary cell wall. The primary wall is deposited extracellularly when cells expand and it consists mainly of cellulose fibrils. These fibrils are embedded in a matrix of structural proteins and polysaccharides, classified in two major groups, the hemicelluloses and the pectins. The expanding cell wall contains approximately 30 % of each polysaccharide class and only 1-5 % of proteins (dry weight), noteworthy to mention that strong deviations are common between plant species [Cosgrove1997]. A water content of 75-80 % in the cell wall suggests the formation of a hydrogel in the extracellular apoplast [Cosgrove1997]. The secondary cell wall is deposited after the development is finished and consists mainly of cellulose enriched with lignin [Brett1996]. The single cell wall compounds are discussed in the following part.

1.4.1 Cellulose

The cellulose fibres are synthesized directly in the extracellular matrix by terminal complexes containing the cellulose synthases. Whereas terminal complexes were first found in freeze fractures of green algae [Kiermayer1979], [Giddings1980] and a little while later after in higher plants [Brown1996], but it was not until the late 1990s that the first cellulose synthase was sequenced [Arioli1998]. The cellulose fibres are assembled to (1,4)-linked β-glucan chains by the plant cellulose synthase protein (CESA) with a size of 3-5 nm width and several microns in length [Cosgrove2005].

Further genes were identified, which are necessary for correct cellulose assembly such as a sterol glycoside [Peng2002] and an endonuclease illustrating the complexity of the cellulose synthesis process [Nicol1998]. Cellulose is organized mainly in crystalline fibers, but can occur in less organized amorphous regions, whereby non-covalent bonds are responsible for their structural arrangement and their extraordinary properties [Cosgrove1997]. Furthermore, cellulose is highly resistant towards chemical and enzymatic treatment.

1.4.2 Hemicelluloses

Hemicelluloses are subjected to a group of heterogeneous, non-crystalline glucans ((1,4)-β-D-glucan backbone) with the ability to bind tightly the cellulose fibrils in parallel orientation to the cellulose fibres [Morikawa1978], [Hayashi1989], preventing their self-aggregation [Cosgrove1997]. The site of synthesis is the Golgi apparatus producing highly branched molecules. According to current nomenclature, xyloglucans are also

referred to as KOH soluble fraction, which is based on their chemical extraction method and gives a more defined fraction. Hemicelluloses can be classified in the two most abundant groups: The xyloglucans (predominant linked to xylose, fucose, glucose and galactose) and the arabinoglucans [Cosgrove1997]. Furthermore, xylanes, arabinoxylans and mannans belong to the hemicelluloses family [Carpita1993].

1.4.3 Pectins

Pectins are a very heterogeneous group of characteristic acidic cell wall polysaccharides, commonly solubilized with hot water treatment or calcium chelators. The multiple functions include an influence on the cell wall porosity and strength, pH balance and mediator molecules in case of pathogen attack [Carpita1993].

Homogalacturonan (HG), rhamnogalacturonan I (RGI) and rhamnogalacturonan II (RGII) are the most abundant species in dicotyledonous plants.

HG's consist of a (1,4)-β-galacturonic acid with occasional rhamnosyl chains. The carboxyl residues of HG's are often methyl-esterified. RGI's have basic units of (1,2)-α-L-rhamnosyl-(1,2)-β-D-galacturonyl disaccharides and arabinans with arabinogalactans as side residues. RGII's are the most variable and complex group of the pectins, which can form dimers through borate ester linkages (the diversity of pectins was reviewed a few times [Carpita1993], [Willats2001], [Vincken2003]). Covalent-linkages were suggested for pectins [Vincken2003] with RGI as backbone polymer and other pectin types as side-branches. HG's build additional ion bridges via calcium and carboxylate groups and borate-dimers [Cosgrove2005]. The calcium bridges can be removed by methyl esterification of the carboxyl groups such when the ionic binding position is blocked and the negative charge is removed. In the expanding cell wall approximately 75 % of the HG's are methyl-esterified [Cosgrove2005]. Pectin methyl esterases (PME) catalyze the hydrolysis reaction to eliminate the methyl group from the pectin residues [Micheli2001].

The biosynthesis of pectins takes place in the endoplasmic reticulum and Golgi apparatus, where they are methyl-esterified and substituted in the *trans* Golgi cisternae [Goldberg1996]. The pectins are delivered in a methyl-esterified state to the cell wall and the modification is subsequently removed via PME activity after the cell expansion process has ceased [Micheli2001].

1.4.4 Cell Wall Proteins

1.4.4.1 Proteins

The cell wall includes diverse groups of structural glycoproteins that are classified on the basis of their amino acid composition [Showalter1993]. The main families are hydroxyproline-rich glycoproteins (HRGP) such as extensins, arabinogalactan proteins (AGP), glycin-rich proteins (GRP) and proline-rich proteins (PRP). Various of these proteins share some properties, including a high glycosylation grade and repetitive elements, which make them insoluble in the cell wall [Cosgrove1997]. Various AGP's are soluble and therefore an exception of the typical cell wall proteins [Fincher1983], [Showalter1993], but others contain lipid tails which are incoperated into the plamsa membrane [Albersheim2011]. Their abundance is highly variable regarding plant species and cell types. The function of the proteins is manifold due to their ability to bind pectins. This includes cell-cell adhesion, development and pathogen defense or they serve as shuttle molecule to keep otherwise insoluble polymers soluble [Gibeaut1991].

1.4.4.2 Enzymes

Enzymes are important compounds for the function of the cell wall giving them a "metabolic activity" [Cosgrove1997]. Many enzymes were identified with polysaccharide modifying function such as glucanases, pectinases and xylosidases, chitinases (defense) or expansins acting in cell wall expansion allowing a pH-dependent wall-loosening.

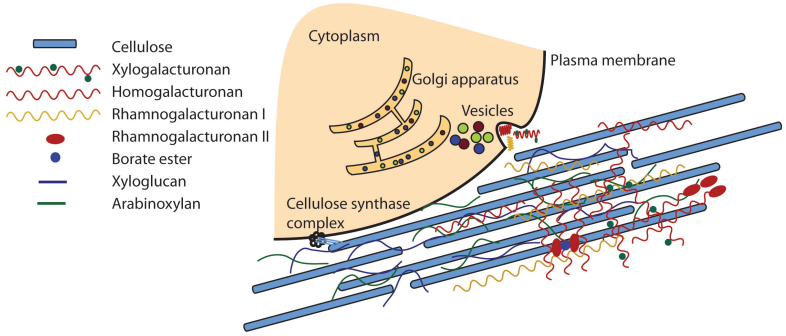


Figure 1.4: Model of the primary cell wall with cellulose fibres synthesized by a multi-protein complex (cellulose synthase complex). Hemicelluloses and pectins are delivered in vesicles to the cell wall after their synthesis in the Golgi apparatus. Hemicelluloses are attached to both, the cellulose fibres (left part), whereby the pectins are suggested to be linked to each other [Vincken2003], and to hemicelluloses [Rizk2000]. Some of them are able to bind cellulose fibrils [Zykwinska2005]. Figure has been modified from [Cosgrove2005].

A recent model of the cell wall structure, including the previously discussed major groups of polysaccharides, is shown in Figure 1.4. In spite of increasing knowledge on how plants build their "organs", many things are lacking, such as the full understanding of the cell wall synthesis.

1.4.4.3 The Plant Cell Wall - Calcium Homeostasis

Calcium plays a major role in diverse cell biological processes in organisms. As already described above, calcium is important for cross-linking of the pectin network in the plant cell wall. Therefore, the calcium content has to be tightly controlled intracellularly and extracellularly in each cell. The calcium efflux on the plasma membrane is achieved by Ca^{2+} ATPase and by H^+/Ca^{2+} antiporters [Briskin1990]. The first Ca^{2+} ATPase was cloned in 2000 [Bonza2000] and a N-terminal calmodulin binding sequence was identified. Calmodulin (calcium binding protein) is known to regulate the activity of Ca^{2+} ATPase in the plasma membrane. The main storage for calcium is the plant vacuole and the calcium concentration in the apoplast is controlled via the transport of calcium to the vacuole. This was demonstrated by a double knock out mutant of H^+/Ca^{2+} antiporter localized in the tonoplast, which accumulated 30 % more free calcium in the apoplast and had a reduced calcium content in the vacuole of mesophyll cells in *Arabidopsis* leaves. Changes in the calcium homeostasis have an impact on diverse processes such as reduced cell wall extensibility, stomatal aperture, transpiration, CO_2 assimilation, leaf growth rate, cell wall protein expression, extensin and cellulose synthase demonstrating the impact of calcium in cell biological processes [Conn2011].

1.5 Biotechnology - Plants as Bioreactors for Protein Biosynthesis

1.5.1 The Advantages of Plants as Bioreactors

Recently, broad spectra of organisms have been used for recombinant protein biosynthesis including structural-, biopharmaceutical-, industrial- or food processing proteins. The plant biomass is a potential source of compounds for industrial and biotechnological approaches. Each system has advantages as well as disadvantages. Originally, bacteria, funghi, animal cells and plants were used as expression systems. Bacteria and funghi are more simple organisms for protein biosynthesis but require more initial capital for the

fermentation equipment and are limited regarding posttranslational protein modification such as glycosylation [Montagnani2011]. This fact may result in protein misfolding or reduced protein solubility [Hood2002]. Animal systems are sufficient for correct post-translational modifications but bear the drawback of being very costly, time consuming and the yield of protein is often very low [Hood2002]. The plant as a bioreactor has a few advantages for protein synthesis such as low costs for protein synthesis and protein stability during the storage such as in seeds. One limiting factor regarding posttranslational modification is a different setup in glycosylation pattern, which is not always identical to animals patterns [Ma2003]. It can be related to the exchange of α-(1,3) fucose and β-(1,2) xylose residues against galactose and sialic acid residues [Ma2003]. To overcome this limitation, different strategies are used such as coexpression of glycosylating enzymes in plants, which results in a human glycosylation pattern [Bakker2001].

Plant protein production is easy to scale up and plant material can often be added directly without large purification steps [Hood2002]. The successful expression was already shown for a variety of proteins with different physical properties. The cost factor for production depends strongly on the protein amount which can be produced by the organisms such that the plants have to express proteins at least in the range of 0.01 to 0.1 % of the tissue's weight. Plants have the advantage to offer different promotors and targeting sequences for a large scale protein biosynthesis in specific plant tissue and are generally safe to human pathogens which is important for food and pharmaceutical production [Hood2002].

System	Overall cost	Production timescale	Scale-up capacity	Product quality	Glycosylation	Contamination risks	Storage cost
Bacteria	Low	Short	High	Low	None	Endotoxins	Moderate
Yeast	Medium	Medium	High	Medium	Incorrect	Low risk	Moderate
Mammalian cell culture	High	Long	Very low	Very high	Correct	Viruses, prions and oncogenic DNA	Expensive
Transgenic animals	High	Very long	Low	Very high	Correct	Viruses, prions and oncogenic DNA	Expensive
Plant cell cultures	Medium	Medium	Medium	High	Minor differences	Low risk	Moderate
Transgenic plants	Very low	Long	Very high	High	Minor differences	Low risk	Inexpensive

Figure 1.5: Comparison of different protein expression systems for human health pharmaceuticals regarding costs, time scale for production, scale up capacity, product quality, glycosylation and contamination risks [Ma2003]. It illustrates the potential of plants as a bioreactor for protein biosynthesis, whereby traditional expression systems have limitations regarding costs, upscaling and safety [Ma2003].

A comparison of different expression systems is given in Figure 1.5 including costs,

timescale for production, scale up capacity, product quality, glycosylation, contamination risks and storage costs [Ma2003].

Therefore, plant expression systems turn out to be cost saving and low risk systems for human health products. The reduced costs are reasons for many companies to be interested in the production of plant proteins for the recent market such as Monsanto, Epicyte Pharmaceutical, Large Scale Biology Crop, Crop Tech Corp., Phytomedics Inc., CollPlant and Prodigene, only to mention a few examples [Ma2003].

1.5.2 Strategies and Examples for Xenogenic Protein Biosynthesis in Plants

A versatile route toward xenogenic (foreign) protein synthesis is the usage of *Agrobacterium tumefaciens* for plant transformation [Gelvin2005], [Hood2002]. The constructs include a promotor region, a targeting sequence, the gene of interest and a terminator sequence. For facilitated selection, a selection marker is co-expressed with the gene of interest. Different promotors can be used for tissue-specific protein targeting including apoplast, vacuole, endoplasmic reticulum, plastids and seeds. Furthermore, the codon usage must be considered in the different expression systems. The ability to synthesize xenogenic proteins in plants depends on various factors, such as transcription, translation, targeting and the accumulation of the protein in the plant [Hood2004]. It is noteworthy to mention that the protein stability and accumulation also depends on the microenvironment such as pH, temperature, protein- and salt content and has to be evaluated empirically for each protein. There are quite a few examples where the subcellular expression was found in unpredicted compartments such as for the LT-B protein described in Chikwamba et al., where the protein was targeted to the maize plant cell wall but was observed in the seed starch granules [Chikwamba2003]. How to enhance xenogenic protein synthesis in plants regarding different cell biological process was recently reviewed [Streatfield2007].

Due to the broad spectrum of applications, a lot of proteins have already been synthesized in plants. The importance of the plant cell wall polymers as targets for protein synthesis is shown by the interest of many industry branches, which depend on cell wall polymers. Paper-, bioethanol-, food- or pharmaceutical industries depend on an inexpensive and renewable source of fermentable sugar [Ziegler2000]. A lot of effort is put into the development of new strategies to simplify the cell wall degradation processes.

For example, one limiting factor for the upscaling process of the cell wall polymers (cellulose, hemicellulose) for bioethanol production is a process called "saccharification", which means the subsequent degradation to simple sugars for ethanol production. Additional modifications of the cell wall such as lignification (lignocellulose) increase the stability of the wall. A pretreatment is necessary to obtain the sugar monomers, which can be used for fermentation processes in bacteria, whereby the quality of the pretreatment strongly influences the further saccharification steps. Most investigated techniques are based on thermochemical processes (reviewed in [Wyman2005]), but also hydrothermal pretreatments were investigated successfully [Kristensen2008]. Nevertheless, recent approaches include breeding and transgenic modification of plants to improve the properties of plant cell wall [Pauly2010].

Enlarged rosette leaves in *Arabidopsis thaliana* with a decrease in xyloglucan-cellulose cross linkages were obtained due to the constitutive expression of a poplar cellulase [Park2003]. Moreover, decrease in xyloglucan content was achieved by double knock out mutants of enzymes involved in the xyloglucan biosynthesis [Cavalier2008]. The expression of a thermostable cellulase from *Acidothermus cellulolyticus* by plant cell protoplast and transgenic *Arabidopsis thaliana* plants realized facilitated bioconversion. The high enzyme activity temperature of the cellulase at 81 °C allows the transgenic plants to grow normally at ambient temperature and has low influence on cellulose deposition in the mother plants [Ziegler2000].

Furthermore, structural proteins such as spider silk were expressed by tobacco, potatoe [Scheller2001] and *Arabidopsis* [Yang2005]. Collagen I was coexpressed with proteins necessary for posttranslational modification in the plant vacuole [Stein2009], [Xu2011]. An overview of vaccines produced in transgenic plants for animal or human targets is given by Hood and Jilka [Hood1999]. Besides the numerous examples for heterologous expressed proteins, only one example in the context of biomineralization and plant biotechnology was recently published [Nakata2012]. Here, two oxalic acid biosynthetic genes, *obcA* and *obcB*, from the oxalate-secreting phytopathogen *Burkholderia glumae*, were inserted into the *Arabidopsis* genome, not only achieved the production of measurable amounts of oxalate but also allowed the formation of calcium oxalate crystals [Nakata2012].

As mentioned before, the xenogenic expression of proteins in plants was successful for various proteins, although there are no "rules" for the success of expression and it has to be investigated empirically for each protein. In the next section, it will be shown, how proteins are secreted in the cell and which cellular processes are available for incorrectly

processed proteins.

1.5.3 Plant Secretory Pathway

The mechanism of the protein synthesis and transport to the apoplast is complex and it is highly essential for various biological processes such as cell wall formation during development, root hair tip growth but also immune response [Meyer2009], [Kwon2008]. The site of synthesis of secretory proteins in plant cells takes place in the rough endoplasmic reticulum (RER), where mRNA is used for protein synthesis (translation). Secretory proteins contain a signal peptide sequence (SP), which was shown to have hydrophobic amino acids.
During translation, a signal recognition peptide (SRP), consisting of protein and RNA, binds to the SP and the ribosome and recognizes a SRP receptor in the RER membrane associated to a translocon [Albersheim2011]. The synthesized peptide chain is elongated into the translocon pore and the SP is released by cleavage due to a peptidase enzyme on the membrane. Sugar residues are assembled and attached to the proteins (N-linked glycoprotein to an asparagin residue), which is transferred via vesicles to the *cis* Golgi apparatus. The Golgi apparatus consists of different zones (*cis, medial, trans* cisternae and *trans* Golgi network) with specific functions. The cisternae are linked via proteins for their stability [Albersheim2011]. In contrast to the single aggregated animal Golgi apparatus, plant cells contain more than one hundred single Golgi stacks in the cytoplasm, which are connected to the cytoskeleton and can be moved in the cell via actin filaments. Through subsequent modification of the proteins in the Golgi compartments, sugar residues are removed (such as mannose) and other sugars are attached. In addition, the O-linked glycosylation takes place in the Golgi apparatus (glycosylation of amino acids with -OH group) [Albersheim2011].
After "ripening" of the proteins, they can be secreted via the *trans* Golgi network. Furthermore, sugar residues are part of the recognition machinery for protein receptors and can protect themselves against degradation processes accomplished by proteases. The secretory process is accomplished with vesicles, which fuse to the cell plasma membrane and release the proteins. Specific coating proteins facilitate the directed protein transport in vesicles through the secretory pathway such as COPII (ER to Golgi) and COPI (Golgi to ER) [Albersheim2011]. Clathrins are necessary for the recycling transport of proteins from the plasma membrane to the cell, where these vesicles remove their coating and fuse to those called prevacuolar compartments delivering the proteins to lytic vacuoles [Albersheim2011].

In case of the immune response, the SNARE-pathogen vesicles induced protein complex (SNARE complexes, N-ethylmaleimide-sensitive factor attachment protein receptor) target the location of pathogen attack [Kwon2008]. Furthermore, plasma membrane receptors, playing a major role in signalling processes, are delivered by secretory pathway due to the interaction with various proteins such as receptor-accessory proteins [Popescu2012]. During the maturing of the proteins, a quality control is important to ensure that only correctly processed proteins leave the endomembrane system such as correct glycosylation, folding, disulfid-bridges and assembly. The major control systems are described in the next section.

1.5.4 Misfolded Protein Degradation System

The main endoplasmic reticulum (ER) associated degradation system (ERAD) is known from yeast and mammalian cells. ERAD eliminates incorrectly matured proteins in the ER and prevents their delivery to the Golgi apparatus. The glycosylation (N-glycosylation) of proteins plays a major role in the ERAD system, which acts as signalling for the calnexin/calreticulin cycle. Both proteins recognize the glycosyl residues, promote folding and prevent early aggregation of the proteins. Correctly folded proteins are released to the Golgi apparatus, whereby protein misfolding leads to a processing of the glycosylation and this modification is recognized by the ERAD system such that the proteins are degradated in the cytosolic proteasome.

The plant degradation system is not well understood, which was reviewed recently [Huettner2012]. Different ERAD target proteins are identified and proteasome dependent and independent (such as vacuolar) protein degradation systems pathways are known. In *Arabidopsis*, one target of ERAD is the BRASSINOSTEROID INSENSITIVE receptor kinase (BRI1), which is a plasma membrane protein and important for the brassinolid pathway [Huettner2012]. A defect ERAD system leads to impaired growth of the plant [Huettner2012]. If misfolded proteins are accumulated, the so-called unfolded protein response (UPR) is activated [Malhotra2007], [Urade2007], [Vitale2008]. Proteins upregulated during UPR are reviewed and can lead to diverse cell biological responses such as protein degradation, transcriptional changes and apoptosis [Urade2007].

It has to be taken into account, that xenogenic protein expression can lead to the activation of degradation systems due to protein overexpression and accumulation or incorrect processing in the ER or Golgi apparatus.

1.6 Objectives of the work

The aim of this work was to improve our understanding of the interactions between calcium carbonate minerals and recombinantly produced biomineralization proteins for the targeted synthesis of artificial biominerals. One major bottleneck regarding such synthetic approaches is the limited availability of native biomineralization proteins. In order to identify new strategies to overcome this limitation, fusion proteins of perlucin, involved in the biomineralization of the mollusc shell *Haliotis laevigata* were selected as the test system for expression in a bacterial host for biotechnological approaches.

Perlucin was selected as one of the most promising candidates for promoting calcium carbonate formation (section 1.1.4.1). The aim was to establish an easy purification procedure under native conditions. The strategy was to produce different perlucin variants, and to purify and characterize them with respect to their properties such as solubility and agglomeration behaviour using various biochemical methods. One crucial point was the development of suitable calcium carbonate precipitation assays in order to validate the activity of the proteins and study their potential influence on calcium carbonate formation. Another goal was to characterize the artificially synthesized biominerals with respect to their structural properties using various microscopic and spectroscopic techniques. The use of biotechnologically produced proteins for biocomposite formation bears tremendous possibilities with a potential impact on the production of high performance materials. Besides this, it may help to study and understand fundamental concepts of biomineralization.

The second part of this thesis targeted the expression of non-plant biomineralization proteins in the plant organisms (*Arabidopsis thaliana* and *Nicotiana benthamiana*), which is a very recent strategy and as such not described in the literature so far. The particular aim here was to investigate the feasibility of the foreign gene expression with respect to opening new routes toward hierarchical composite materials produced directly by organisms as a long-term goal.

The biomineralization proteins perlucin (mollusc), N16N (mollusc) and ovocleidin-17 (eggshell) served as the first case studies for the plant expression system (section 1.1.4.2 and 1.1.4.3). The first objective was to identify and select suitable signal peptide sequences for the delivery of proteins to the plant cell wall and to design the cloning strategy. The second objective was to accomplish the transformation of the plant model organisms and to analyse the plant material regarding phenotype, gene- and protein expression level. In order to achieve this goal it was required to use, adapt and estab-

lish suitable molecular biological, biochemical and microscopic techniques. In general, plants are very useful to serve as bioreactors for foreign protein expression. On the other hand, the complexity of the plant organism and in particular the crucial role of the cell wall in many physiological pathways of the entire plant metabolism has to be taken into account. Therefore, it was not possible to predict whether or not the expression of biomineralization proteins in plant cell walls would be successful. In particular, it can not be predicted how a plant will respond once biomineralization proteins such as perlucin interfere with calcium signalling and plant specific processes. Therefore, this second part of the thesis covers a high risk project.

2 Material and Methods

All materials, equipment and chemicals were prepared according to molecular biological standards and are not further mentioned. Generally used buffers were prepared according to [Sambrook2001].

2.1 Materials

2.1.1 Plant Materials

- *Nicotiana benthamiana* (tobacco)
- *Arabidopsis thaliana*

2.1.2 Bacterial Strands

2.1.2.1 Escherichia coli

- *ccdB* One Shot Survival T1-Phage Resistance Cells (Invitrogen)
  Reproduction of gateway vectors without insert and with *ccdB* gene.
  F$^-$*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*Z ΔM15Δ*lac*X74 *rec*A1 *ara*Δ139Δ (*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G *fhu*A::*IS2*
- One Shot TOP10 Cells (Invitrogen)
  Reproduction of gateway vectors with integrated gene of interest.
  F$^-$*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G λ-
- XL1-Blue Cells (Stratagene)
  Reproduction of vectors for recombinant protein expression.
  recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' *proAB lacI*^{q}*ZΔM15* Tn*10* (*Tet*^{r})]

2.1.2.2 Agrobacterium tumefaciens

C58C1 Derivat GV2260/pGV2260; Rif^R, Cn^R

2.1.3 Plasmids

2.1.3.1 Bacterial Expression Plasmids

pENTR/D-TOPO (Invitrogen) All synthetic genes were subcloned and delivered in the pENTR/D-TOPO vector system, Figure 2.1. All sequences are listed in the appendix section 5.1.

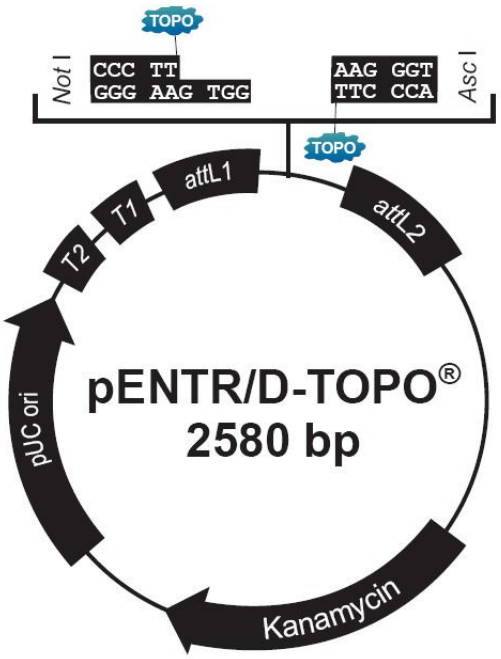


Figure 2.1: pENTR/D-TOPO is a gateway entry vector. The vector contains the gateway recombination sites *attL1* and *attL2* and a 5- overlap for easy cloning procedure of DNA fragments. The vector has a kanamycin resistance gene and an origin of replication (pUC ori) for the selection and reproduction in *E. coli*. The TOPO cloning site is flanked by the two restriction enzymes *Not*I and *Asc*I

.

pDONR207 The pDONR207 vector was used to change the kanamycin resistance gene in the pENTR/D-TOPO plasmid to a gentamycin resistance gene (see for map Figure 2.2). The resistance genes were cut from each vector and the gentamycin resistance gene was ligated into the pENTR/D-TOPO vector.

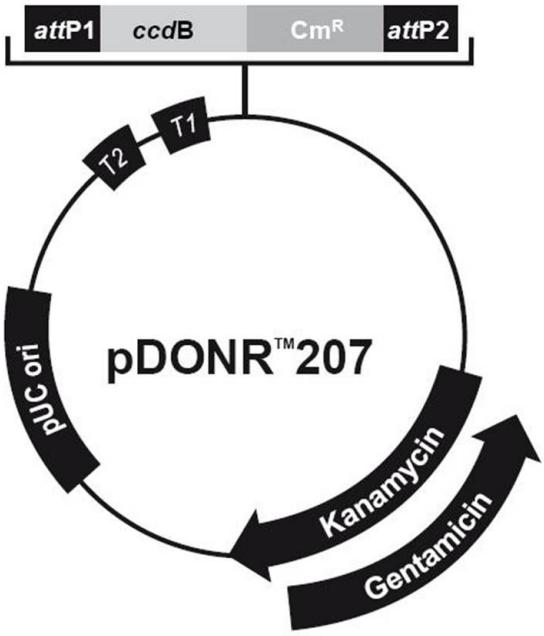


Figure 2.2: pDONR207 is a gateway entry vector containing the recombination sites *attL1* and *attL2*. The vector has a gentamycin resistance gene, which allows a selective reproduction of the plasmid in *E. coli*. The vector was used to exchange the gentamycin resistance gene from this vector against the kanamycin resistance gene in the pENTR/D-TOPO vector.

pQE30/1 (Qiagen) The pQE vector system from Qiagen is optimized for the expression and the purification of recombinant proteins with fusion to a HIS_6 tag (Figure 2.3). The vector system was used for the expression of biomineralization proteins in bacteria.

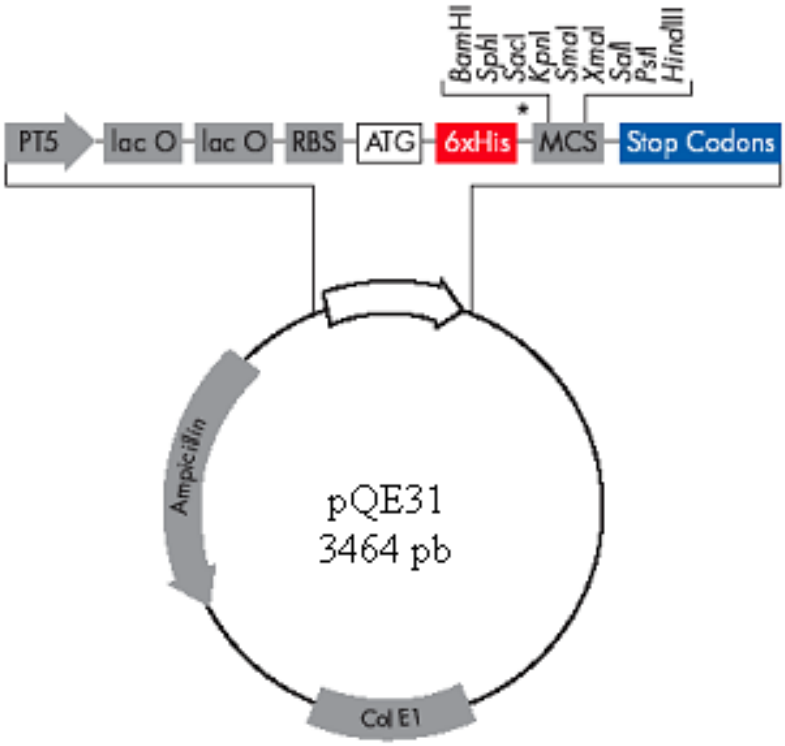


Figure 2.3: pQE31 and pQE30 (not shown) contain a multiple cloning site (MCS) for cloning procedures. The HIS_6 tag allows an one step purification process for the protein of interest. The lac operon allows an induction of protein expression by addition of IPTG to the cell culture. The vector has an ampicillin resistance gene for the selection of the transformed bacteria in *E. coli*. The pQE30 and 31 vectors differ in two additional base pairs (AC in pQE31) in the multiple cloning site (MCS) for an in frame cloning of the gene of interest.

2.1.3.2 Plant Expression Plasmids

pMDC83 The pMDC83 is a binary gateway destination vector for the constitutive expression of proteins in the plant cell [Curtis2003]. The plasmid map is shown in Figure 2.4.

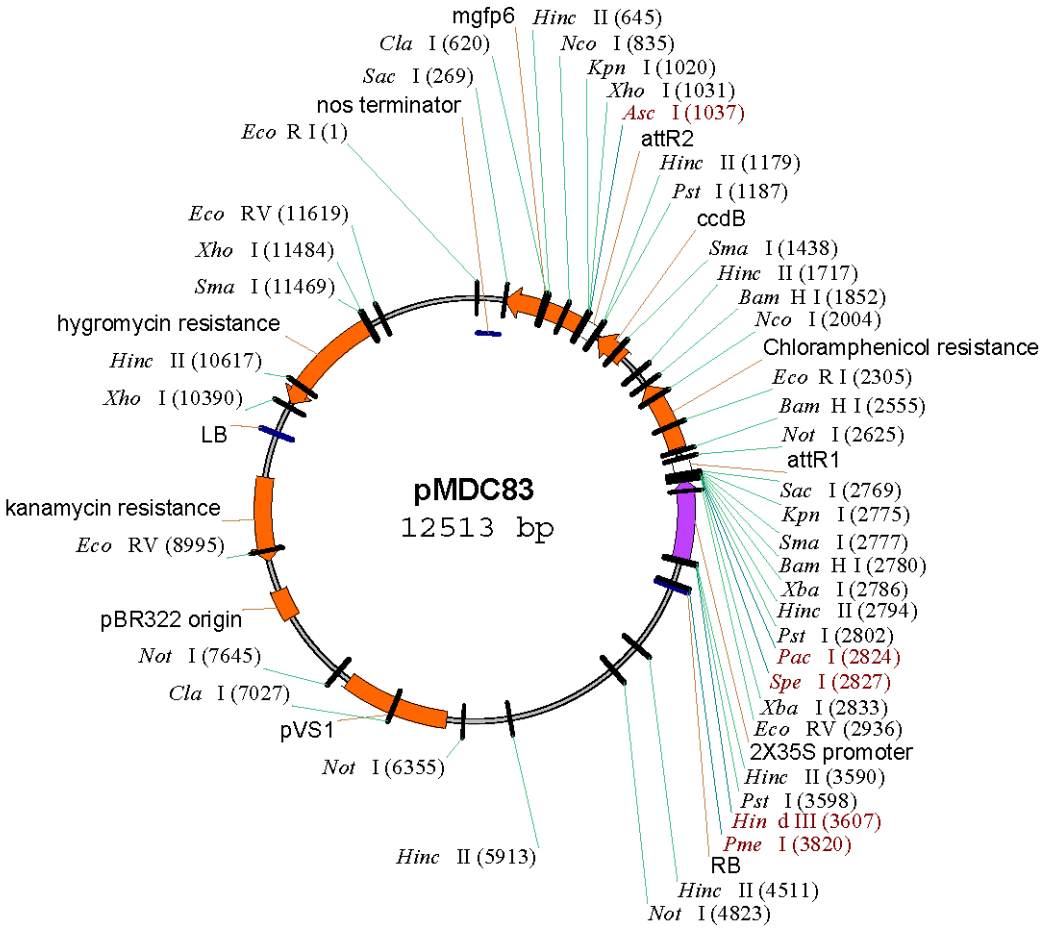


Figure 2.4: The pMDC83 is a gateway plant expression vector for the constitutive expression of the gene of interest under the control of the 2xCaMV35s promotor. The plasmid contains the recombination sites *attR1* and *attR2*. The gene of interest is fused to both, a HIS_6 and a *GFP* gene sequence for easy detection of the protein. The vector contains two selection marker genes, kanamycin for the selection in bacteria and hygromycin for the selection of positive transformed plants.

pALLIGATOR2 The pALLIGATOR2 is a gateway plant expression vector for the constitutive expression of the gene of interest [Bensmihen2004]. The vector map is shown in Figure 2.5.

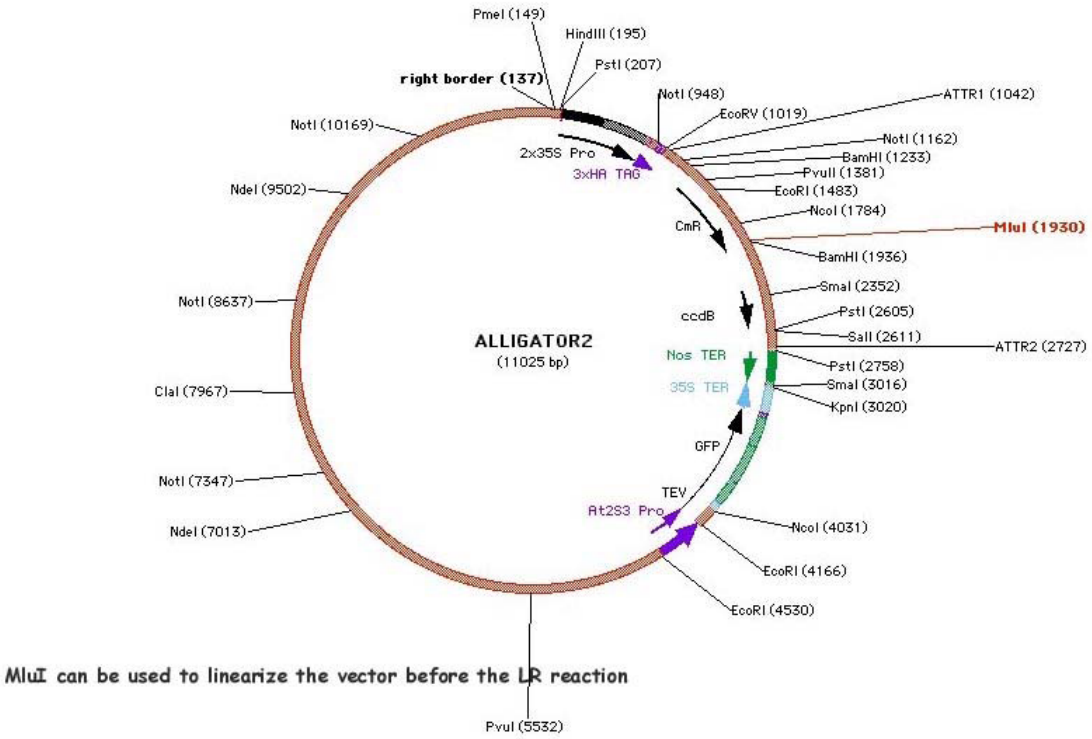


Figure 2.5: pALLIGATOR2 is a gateway destination vector. It contains the recombination sites *attR1* and *attR2* and a 2xCaMV35s promotor for constitutive protein expression in the plant cell. The gene of interest is fused to a HA_3 gene sequence for facilitated protein detection. The vector can be selected with its spectinomycin resistance gene in bacteria. The vector contains an additional *GFP* gene, which is under control of an embryonal seed promotor. Positively transformed seeds are fluorescent and can be distinguished from non-fluorescent wild type seeds.

pMDC30 and 32 Constructs were prepared for both vector systems, but have not been analyzed so far. Both plasmids have been slightly modified from the pMDC83 vector [Curtis2003].

The pMDC30 vector contains an inducible promotor for directed gene expression. The pMDC32 vector contains a 2xCaMV35s promotor but no *GFP* gene sequence (vector

maps are not shown). The vector maps and expression cassettes are described in Curtis 2003 ([Curtis2003]).

2.1.4 Kits

- GenElute^{TM} HP Plasmid Miniprep Kit, Sigma Aldrich, Munich, Germany
- NucleoSpin^{R} ExtraktII, Macherey and Nagel, Düren, Germany
- Spectrum^{TM} Plant Total RNA Kit, Sigma Aldrich, Munich, Germany
- Extract-N-Amp^{TM} Plant Kit, Sigma Aldrich, Munich, Germany
- RevertAid^{TM} First Strand cDNA Synthesis Kit, Fermentas, St. Leon-Rot, Germany

2.1.5 Chemicals

Acetic acid	Roth, Karlsruhe, Germany
Acetosyringon	Roth, Karlsruhe, Germany
Agarose	Roth, Karlsruhe, Germany
Albumin fraction V, protease free BSA	Roth, Karlsruhe, Germany
Ammonium persulphate	Sigma Aldrich, Munich, Germany
L-Arginine	Fluka, Munich, Germany
Bacto agar	Duchefa, Haarlem, The Netherlands
Bacto peptone	Difco VWR, Darmstadt, Germany
Bacto yeast extract	Difco VWR, Darmstadt, Germany
5-Bromo-4-chloro-3-indolyphosphate p-toluidine salt	Roth, Karlsruhe, Germany
Bromphenol blue sodium salt	Merck, Darmstadt, Germany
Coomassie brillant blue R 250	Merck, Darmstadt, Germany
Dimethyl sulfoxide (DMSO)	Roth, Karlsruhe, Germany
Dipotassium hydrogen phosphate (KH_2PO_4)	Fluka, Munich, Germany
Ethidium bromide 10 mg/ml	Bio-Rad, Munich, Germany
Ethylenediaminetetraacetic acid (EDTA)	Roth, Karlsruhe, Germany
FITC-Concanavalin A	Sigma Aldrich, Munich, Germany
Glucose	Roth, Karlsruhe, Germany
Glycerol (x %)	Fluka, Munich, Germany
Glycine	Roth, Karlsruhe, Germany
Hydrochloric acid	Roth, Karlsruhe, Germany

Imidazole	Roth, Karlsruhe, Germany
Isopropyl-β-D-thiogalactopyranosid (IPTG)	Roth, Karlsruhe, Germany
2-Mercaptoethanol	Roth, Karlsruhe, Germany
Methanol	Roth, Karlsruhe, Germany
Micro agar	Duchefa, Haarlem, The Netherlands
Milk powder	Roth, Karlsruhe, Germany
Murashige and skoog medium	Duchefa, Haarlem, The Netherlands
Ni-NTA-agarose	Qiagen, Hilden, Germany
Nitroblue tetrazolium salt (NBT)	Roth, Karlsruhe, Germany
N-N-Dimethyl formamide	Sigma Aldrich, Munich, Germany
Plant agar	Duchefa, Haarlem, The Netherlands
Ponceau	Merck, Darmstadt, Germany
Potassium chloride	Sigma Aldrich, Munich, Germany
Potassium hydrogen phosphate ($KHPO_4$)	Sigma Aldrich , Munich, Germany
Potassium hydroxide	Roth, Karlsruhe, Germany
Potassium nitrate	Roth, Karlsruhe, Germany
Potassium phosphate	Fluka, Munich, Germany
Silwet gold	Spiess-Urania, Hamburg, Germany
Sodium chloride	Fluka, Munich, Germany
Sodium dodecyl sulfate (SDS)	Roth, Karlsruhe, Germany
Sodium hydrogen carbonate	Roth, Karlsruhe, Germany
Sodium hydroxide	Roth, Karlsruhe, Germany
Sodium hydrogen phosphate (Na_2HPO_4)	Fluka, Munich, Germany
Sucrose	Fluka, Munich, Germany
Tetramethylethylendiamine	Sigma Aldrich, Munich, Germany
Tris(hydroxymethyl)-aminomethan (Tris)	Sigma Aldrich, Munich, Germany
Triton-X-100	Roth, Karlsruhe, Germany
Tween 20	Sigma Aldrich, Munich, Germany

2.1.6 Antibiotics

All utilized antibiotics are listed in Table 2.1 with their stock and working concentrations.

Table 2.1: Overview of the antibiotics used for the selection of transformed bacteria and plants.

Antibiotic	Stock concentration	Working concentration	Supplier
	[mg/ml]	[μg/ml]	
Ampicillin	50	50-100	Roth
Carbenicillin	34	34	Roth
Gentamycin sulphate	10	10	Roth
Kanamycin sulphate	10	10	Roth
Rifampicin	50	50	Sigma Aldrich
Spectinomycin	50	50	Sigma Aldrich

2.1.7 Enzymes

LR Clonase	Invitrogen, Darmstadt, Germany
Polymerase, Dream Taq	Fermentas, St. Leon-Rot, Germany
Polymerase, TaKaRa Mix	TaKaRa, Mobitec, Göttingen, Germany
Polymerase, Phusion	Finnzymes, Schwerte, Germany
Restriction Enzyme	Fermentas and New England Biolabs
Lysozyme	Fluka, Munich, Germany

2.1.8 Antibodies

All antibodies used in Western blotting experiments are listed in Table 2.2.

Table 2.2: Overview of the antibodies used in protein detection assays.

Antibody	Supplier	Order number
Alkaline phosphatase conj. (goat anti mouse)	Dianova	115-055-062
Peroxidase conj. (rabbit anti mouse)	Sigma Aldrich	A9044-2ML
Anti-HIS_6 tag, mouse IgG, monoclonal	Dianova	DIA900
Anit-GFP, mouse IgG, monoclonal	Roche	11814460001

2.1.9 Diverse Materials

Bradford reagent	Sigma Aldrich, Munich, Germany
Centrifugal devices (MWCO 3 kDa)	Pall, Ann Arbor, USA
Dialysis tubes (MWCO 4-6 kDa)	Roth, Karlsruhe, Germany
ECL Western blotting reagent	GE Healthcare, Munich, Germany
Filter membrane (sterile 0.22 μm)	Millipore, Schwalbach, Germany
Microscopic slides	Ibidi, Martiensried, Germany
Mini dialyzer units (MWCO 7-10 kDa)	Fisher Scientific, St.Leon-Rot, Germany
Ni-NTA matrix	Qiagen, Hilden, Germany
PVDF membrane (0.45 μm)	Millipore, Schwalbach, Germany
Superose 12 10/300GL column	GE Healthcare, Munich, Germany
Turf balls (Jiffy7)	Fa. Meyer, Rellingen, Germany
UVette	Eppendorf, Hamburg, Germany
UV star half area plate	Greiner, Frickenhausen, Germany
Vermiculite Kl 2	Fa. Meyer, Frankfurt, Germany

2.1.10 Oligonucleotids

Table 2.3 lists all PCR primer used in the different experiments.

Table 2.3: Primers used for PCR reaction are listed by name, concentration and sequence.

Primer	Stock concentration	DNA sequence
	[pmol/μl]	5- to 3-
3-Perlucin-PstI	50	TGCACTGCAGTTATCTTTGTTGCAGATTGG
5-Perlucin-KpnI	50	CGGGGTACCGGATGTCCTTTGGG
5-Perlucin-KpnI-pQE31	50	CGGGGTACCCGGATGTCCTTTGGG
3-S-pMDC83	50	GGGTACCGAGCTCGAATTATCACAAG
5-pQE31-71	100	GGATAACAATTTCACACAGAATTC
3-pQE31-195	100	AACAAATCCAGATGGAGTTCTG
5-Gm resistance	100	CCGACCCAGCTTTCTTGTA
3-Gm resistance	100	ATGGCTCATAACACCCCTTG
5-EXT3	50	CACCATGGGTTCTCCTATGGCCTCTT
5-PRP4	50	CACCATGCGCATCTTACCTGAACCTA
3-Perlucin	50	TCTTTGTTGCAGATTGGCGTGAAGC
5-mGFP	100	AAAGGGCAGATTGTGTGGAC
3-mGFP	100	GGACGACGGGAACTACAAGA
5-Zat12	100	TCGCATCCTTGTCCCATATGT
3-Zat12	100	TTCAAATTGTCCACCATCCCTAG
5-2x35s	100	GATAGTGGGATTGTGCGTCA
3-2x35s	100	CAAATGCCATCATTGCGATA
5-RT-Perlucin	100	GAGCTTAGACGTGATCTTGG
3-RT-Perlucin	100	AGCGAATCTCTTTGCTGTAG
3-N16N	50	GCACTTTTTATCCCCGTTGTCGTATC
3-OC17	50	CGCCGCCGCTTTGCAAACGAAGGCG
5-EF1B-alpha-g	100	CCGGGACATATGGAGGTAAG
3-EF1B-alpha-g	100	TCCGAACAATACCAGAACTACG
5-EF1B-alpha-c	100	ACTTGTACCAGTTGGTTATGGG
3-EF1B-alpha-c	100	CTGGATGTACTCGTTGTTAGGC

2.1.11 Gen Synthesis and Sequences

All DNA sequences were synthesized by Entelechon GmbH, Regensburg, Germany. The codon usage was optimized for the plant model organisms in order to fulfill the require-

ments for a successful protein expression (the *perlucin* gene sequence was optimized for *Nicotiana benthamiana* and the *Ovocleidin 17 (OC-17)* and *N16N* gene sequences were optimized for *Arabidopsis thaliana).* The signal peptides for the protein delivery to the plant cell wall originated from *Arabidopsis thaliana.* The sequences of the signal peptides (*EXT3* and *PRP4*) were fused to the respective gene sequence encoding each biomineralization protein. The sequences are specified in appendix 5.1.

2.1.12 Sequencing

All sequencing reactions were performed by the company MWG Eurofins Operon.

2.1.13 Solutions and Buffers

If not stated otherwise, all ingredients were dissolved in water (deionized) and the pH was adjusted according to experimental requirements.

MS Medium for Plant Agar Plates	
MS agar	4.3 g/L
Sucrose	1 %
Plant agar	0.8 %

MS agar and the sucrose was dissolved in water and adjusted to pH 6.0. The plant agar was added before autoclaving. Hygromycin was added to reach a final concentration of 25 μg/ml for hygromycin selection of transformed *Arabidopsis thaliana* plants.

Seed Sterilization Solution	
Sodium hypochlorite	6 %
Triton-X-100	0.1 %

Coomassie (1 L)	
Coomassie R 250	0.25 %
Methanol	250 ml
Acetic acid	200 ml

TNK Buffer (200 ml)	
NaCl	0.2 M
KCl	0.2 M
Tris/Cl, pH7.5	10 mM

AP Buffer (200 ml)	
Tris/Cl	2.42 g
NaCl	1.17 g

Tris and NaCl were dissolved in ∼150 ml of water (pre-adjusted to pH 9.8) and combined with 2.03 g of $MgCl_2$ x 6 H_2O in 30 ml of water. The pH was adjusted to 9.5 and filled up to 200 ml.

Developer Solution for Alkaline Phosphatase (freshly prepared)	
NaH_2PO_4	2 mM
2.5 μl	75 mg/ml NBT in 70 % DMF
15 μl	50 mg/ml BrCIP in DMF, anhydrous

The solution was prepared in 5 ml AP buffer prior to use.

Infiltration Medium for Tobacco Plants	
NaH_2PO_4	2 mM
MES buffer	50 mM
Glucose	0.5 %

Following the preparation of the solution, the pH was adjusted to 5.6. Acetosyringon (stock: 200 mM in DMSO) was added prior to use to a final concentration of 100 μM.

Infiltration Medium for *Arabidopsis thaliana* Plants (The Floral Dip)	
$MgCl_2$	10 mM
MES/KOH	10 mM pH 5.2 (KOH)
Sucrose	5 %
Silwet gold	0.05 %

100 μM acetosyringon (stock: 200 mM in DMSO) was added prior to use.

TE Buffer

Tris/Cl, pH 7.4	10 mM
EDTA, pH 8.0	1 mM

TAE Buffer for DNA and RNA Gels

50 x stock solution 1 L	
(working solution 1 x)	
Tris base	242 g
Acetic acid	57.1 ml
EDTA	100 ml (0.5 M, pH 8.0)

Plant DNA Extraction Buffer

Tris/Cl, pH 7.5	200 mM
NaCl	200 mM
EDTA, pH 8.0	25 mM
SDS	0.5 %

Plant Protein Extraction Buffer

Tris/Cl, pH 8.0	750 mM
Glycerol	4 %
β-Mercaptoethanol	100 mM
SDS	2 %
EDTA	0.1 %

E. coli Protein Extraction Buffer

Basic buffer:	
NaH_2PO_4	50 mM
NaCl	300 mM
Add imidazol for:	
Lysis buffer	10 mM
Wash buffer	20 mM
Elution buffer	500 mM

Table 2.4 shows the pipetting scheme for SDS polyacrylamid gel preparation.

Table 2.4: Pipetting scheme for 10 % SDS-PAGE gel preparation.

10 %	1 x stacking gel	1 x separation gel
Tris 1 M	330 μl (pH 6.8)	3.05 ml (pH 8.8)
H_2O Millipore	1.535 ml	1.34 ml
SDS 10 %	28 μl	80 μl
TEMED 5 %	60 μl	165 μl
Acrylamid 30 %	430 μl	2.44 ml
PER 20 mg/ml	110 μl	220 μl

Native Page Sample Buffer 2 x

Tris/Cl, pH 6.8	62.5 mM
Glycerol	40 %
Bromphenol blue	0.01 %

Native Page Running Buffer 10 x

Tris/Cl	250 mM
Glycin	1.92 M

The pH was adjusted to 8.3.

Fixing Solution for Protein Silver Staining

Methanol	50 ml
Acetic acid	12 ml
Formaldehyde (37 %)	50 μl

The solution was filled up to 100 ml with H_2O_{MP} and prepared prior to use.

Sodium Thiosulphate Solution for Protein Silver Staining

$Na_2S_2O_3$	0.02 g

The solution was filled up to 100 ml with H_2O_{MP} and prepared prior to use.

Impregnation Solution for Protein Silver Staining

$AgNO_3$	100 mg
Formaldehyde (37 %)	75 μl

The solution was filled up to 100 ml with H_2O_{MP} and prepared prior to use.

Developer Solution for Protein Silver Staining

Na_2CO_3	6 g
$Na_2S_2O_3$	200 μl (0.2 g/l)
Formaldehyde (37 %)	50 μl

The solution was filled up to 100 ml with H_2O_{MP} and prepared prior to use.

Stopping Solution for Protein Silver Staining

Acetic Acid	5 ml
H_2O_{MP}	95 ml

Precursor Solutions for Inorganic ($CaCO_3$) Precipitation Assays-$CaCl_2$

$CaCl_2$	20 mM
Tris/Cl, pH 8.7	3 mM
Sterile filtration	0.22 μm

Solutions for Inorganic ($CaCO_3$) Precipitation Assays-$NaHCO_3$

$NaHCO_3$	20 mM
Tris/Cl, pH 8.7	3 mM
Sterile filtration	0.22 μm

2.1.14 Molecular Weight Standards

Protein and DNA molecular weight standards used for molecular and biochemical analysis are shown in Figure 2.6.

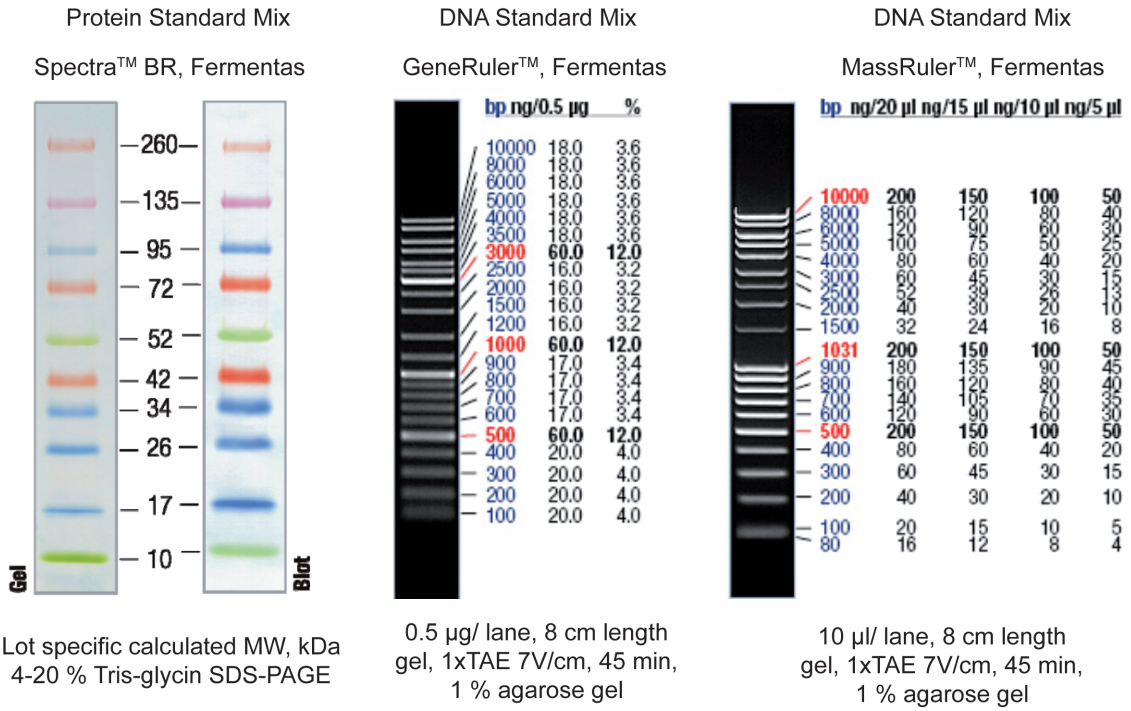


Figure 2.6: Protein and DNA molecular weight standards were used for biochemical and molecular biological analysis.

2.1.15 Equipment

Accurancy balance	LP620S, Sartorius, Göttingen, Germany
Äkta purifier system	GE Healtcare, Munich, Germany
Bioillumination system	FluorchemQ, Biozym, Hessisch Oldendorf, Germany
Blotting apparatur	Mini-Trans-Blot, Bio-Rad, Munich, Germany
Capillary gel electrophoresis	Qiaxcell, Qiagen, Hilden, Germany
Centrifuge	Rotana 460RS, Hettich, Tuttlingen, Germany
Drying oven	UT6060, Heraeus, St.Leon-Rot, Germany
Field-flow fractionation HPLC	Agilent, Böblingen, Germany
Field-flow fractionation AF4	Wyatt, Dernbach, Germany
Freezer -80 °C	New Brunswick, Hamburg, Germany

Gel electrophoresis (DNA, RNA)	HU10, 20, Fisherbrand
Gel electrophoresis (proteins)	Mini-Protean Tetra Cell, Bio-Rad, Munich, Germany
Incubator	Innova 42, New Brunswick, Hamburg, Germany
PCR cycler	Master Cycler Gradient, Eppendorf, Hamburg, Germany
PH electrode	N6000, SI Analytics, Jena, Germany
PH meter	Problab 3000, Schott, SI Analytics, Jena, Germany
Photometer	Biophotometer, Eppendorf, Hamburg, Germany
Plant growth chamber	AR22L, CLF Plant Climatics, Wertingen, Germany
Plant growth chamber	AR36L, CLF Plant Climatics, Wertingen, Germany
Plant growth chamber	AR95-LX, CLF Plant Climatics, Wertingen, Germany
Plant growth chamber	ATC60-Flex, Conviron, Manitoba, Canada
Power supply	Zoom Dual Power, Invitrogen, Darmstadt, Germany
Real-time PCR cycler	Mastercycler Realplex, Eppendorf, Hamburg, Germany
Shaker	Shaker DRS-12, Neolab, Heidelberg, Germany
Spectra Max 190	Molecular Devices, Sunnyvale, USA
Thermomixer	Thermomixer comfort, Eppendorf, Hamburg, Germany
Table centrifuge	Centrifuge 5415 R, Eppendorf, Hamburg, Germany
"Ultra" centrifuge	Avanti J-26 XP, Beckman Coulter, Krefeld, Germany
Waterbath	Memmert, Schwabach, Germany

2.1.16 Software and Web Tools

- Astra, version 5.3.4, for data analysis with the field-flow fractionation system (Wyatt Technology Corporation)
- ClustalW, for sequence alignment analysis (http://www.ebi.ac.uk/Tools/clustalw2/index.html, doi:10.1093/bioinformatics/btm404)
- Primerdesign Tool, for oligonucleotide analysis version 0.4.0 (http://frodo.wi.mit.edu/primer3/)
- Ape Software version 1.17, for vector design (http://www.biology.utah.edu/jorgensen/wayned/ape/)
- Ballview version 1.4.1, for 3-D structure analysis of proteins (http://www.ballview.org/)
- WebFeature 2.0, calculates potential calcium binding sites in a protein model (http://feature.stanford.edu/webfeature/)
- Tair, The Arabidopsis Information Resource Site, for general information (http://www.arabidopsis.org/)

- SignalP 3.0, for signal peptide analysis (http://www.cbs.dtu.dk/services/SignalP/)
- iPSORT, for signal peptide analysis (http://ipsort.hgc.jp/how.html)

2.2 Methods

2.2.1 Molecular Biological Methods

2.2.1.1 Isolation of Genomic DNA from Plant Material

The plant material was ground in 400 μl DNA extraction buffer (200 mM Tris/Cl pH 7.5), 200 mM NaCl, 25 mM EDTA pH 8.0, 0.5 % SDS). After rigorous mixing, 400 μl of isopropyl alcohol was added and the sample was mixed again. The cell debris was pelletized through 5 min of centrifugation with 10,000xg. The supernatant was discarded and the pellet was washed three times with 70 % of ethanol. The pellet was air dried after the last centrifugation step and finally resuspended in a 100 μl TE buffer. Alternatively, DNA was extracted with the Extract-N-Amp^{TM} Plant Kit, according to manufacture guidelines (Sigma Aldrich). The DNA extracts were used for PCR reactions or stored at -20 °C.

2.2.1.2 Isolation of RNA from Plant Material

The plant RNA was extracted with the Spectrum^{TM} Plant Total RNA Kit according the manufacture guidelines (Sigma Aldrich, protocol 4A). For one sample 60 seedlings (6 d old) were ground under liquid nitrogen. Following grinding, 100 mg of the prepared powder was used for the extraction procedure and RNA purification was performed using sterile filter tips for pipetting. Finally, the RNA was eluted in a volume of 50 μl.

2.2.1.3 DNA Quantification

The DNA was quantified on a standard agarose gel in combination with a MassRuler^{TM} (Figure 2.6, right) or quantified optically in a photometer. 5, 10 and 15 μl of MassRuler^{TM} were loaded on a gel together with the DNA for agarose gel quantification. The DNA amount was estimated by the intensity of the DNA bands. For the optical calculation of the DNA concentration, 98 μl of autoclaved H_2O was filled into a UVette (Eppendorf).

After blank measurement 2 μl of DNA was mixed with the 98 μl of H_2O (in the same cuvette) and measured again. The DNA concentration was calculated by software integrated in the photometer. The purity of the DNA and RNA was calculated by the 260 nm to 280 nm ratio and are to be in between 1.8 to 2.0.

RNA Quantification The RNA concentration was measured after extraction from the plant material (section 2.2.1.2). To measure the RNA concentration, 1 μl of the extracted RNA was mixed with 80 μl sterilized H_2O_{MP} in a 96 well UV star half area plate (Greiner). Values at the wavelengths of 260 nm and 280 nm were determined in a plate reader. The ratio of 260 nm to 280 nm was calculated after the substraction of the zero control. All values were multiplied with 2 - half area plate (half of the volume of a standard plate) and multiplied with 40 (corresponding value for single stranded nucleic acid). The ratio are to be in between 1.8 and 2.0 for a good purity of RNA.

2.2.1.4 Preparation of DNA and RNA Agarose Gel

Standard agarose gel electrophoresis for DNA and RNA was performed in 1 x TAE buffer solution (for gel preparation and running buffer). The concentration of agarose varied in the gel depending on the DNA fragment size. RNA was separated by using 0.8 % agarose. Prior to experiment, the equipment was cleaned with H_2O_{MP}. Standard gel running conditions were 75-100 V for 30 to 45 min. The samples were prepared with 1 x loading dye (final). 7 μl of GeneRuler^{TM} (Figure 2.6, center) was loaded on each gel for the size calculation. All gels were stained in an ethidium bromide bath (50 μl of a 10 mg/ml ethidium bromide in 1 liter deionized H_2O) for 15 min. The DNA/RNA was visualised and photographed under UV light with a bioillumination system (Biozyme).

2.2.1.5 Capillary Gel Electrophoresis

Capillary gel electrophoresis was used for high throughput genotyping analysis of transformed plant material. Standard PCR analysis were performed with a screening cartridge and an alignment marker of 15-3,000 bp size in a Qiaxcell system (Qiagen).

2.2.1.6 Real Time Analysis of Plant Material

The quantitative real time analysis (qPCR) enables accurate measurement of gene activity on the basis of the mRNA level in a sample. The first step was the RNA extraction

from the plant material (section 2.2.1.2). The extracted mRNA was transcribed in cDNA and the cDNA was used as template for a PCR to amplify the gene of interest. SYBR green was used for the detection, which will be incoperated into each new double stranded DNA amplified and the amount of DNA strands can be calculated through a DNA standard curve where the copy number of DNA is known. The cycle threshold (Ct) is determined, which is defined by the value upon which the fluorescence signal exceeds the background noise. Before DNAse digestion, the RNA concentration was optimized to 1 μg in 7 μl H_2O (RNAse and DNAse free) and diluted such that each sample contained equal amounts of RNA.

DNAse Digestion Before cDNA synthesis, a DNAse digestion of isolated RNA was performed in order to avoid false positive signals in the PCR reaction from genomic DNA contaminations in the RNA extract. Reagents were used from the RevertAid^{TM} First Strand cDNA Kit. All steps were performed using DNAse and RNAse free tubes, H_2O (RNAse and DNAse free) and filter tips for pipetting to avoid degradation of RNA. In addition, a negative control (non-template control) was prepared without RNA.

Reaction	1 x
RNA	1 μg (in 7 μl H_2O)
DNAse buffer 10 x	1 μl
RNAse inhibitor 10 units	0.1 μl
DNAse	1 μl
Volume	10 μl

The DNAse digestion was performed at 37 °C for 30 min in a PCR cycler.

Inactivation of DNAse and Primer Annealing The next step was the inactivation of the DNAse and the primer annealing for the cDNA synthesis. All solutions were directly added to the tube. Oligo(dt) primers were used for the reaction, which are complementary to the poly-A-tail of eukaryotic mRNA.

Reaction	1 x
25 mM EDTA	1 μg (in 7 μl H_2O)
Oligo$(dt)_{18}$ mix	1 μl
Volume	12 μl

The inactivation of the DNAse and the denaturation of the RNA was performed for 10 min at 65 °C.

cDNA Synthesis Subsequently, the cDNA synthesis was performed in the same tube.

Reaction	1 x
H_2O nuclease free	1 μl
Buffer 5 x	4 μl
RNAse inhibitor 10 units	0.5 μl
dNTP 10 mM each	2 μl
M-MuLV RTase	0.5 μl
Volume	20 μl

The cDNA synthesis was performed for 2 h at 42 °C and the enzyme was inactivated for 10 min at 65 °C. 180 μl of nuclease free H_2O was added to each tube. The obtained cDNA solution was stored at -20 °C. The reaction was diluted 1:10 before the PCR reaction.

qPCR Reaction A PCR was performed for each synthesized cDNA sample and the corresponding gene standards (see for preparation section 2.2.1.6) with two additional negative controls (H_2O control and cDNA synthesis (non-template) control). In addition, a housekeeping gene *EF1Balpha2* was amplified as internal control for intact cDNA and for normalization of the measurements. Primers were used for the amplification of both, a gene sequence without (Efc) and with an intron (Efg) sequence. The latter allowed a correction of the measured values in case of contamination with genomic DNA.

Table 2.5: Master mix for the qPCR reaction.

Master Mix	Amount [μl]
H_2O nuclease free	1
TaKaRa Premix 2 x	10
5-Primer	0.2
3-Primer	0.2
SYBR green	0.1

Table 2.6: Temperature profile for the qPCR reaction.

Step	Temperature [°C]	Time	Repeats
Denaturation	95	2 min	
Denaturation	95	10 sec	
Annealing	58	15 sec	40
Elongation	72	15 sec	
Melting curve	60 - 95 C	20 min	

The optimal concentration for each individual primer pair was analyzed in a primer matrix before the PCR reaction, as described in section 2.2.1.6. For the qPCR reaction 10 μl of the freshly prepared master mix (Table 2.5) and 10 μl of the cDNA, standards and negative controls were pipetted into each well of a 96 well plate. A schematic example of 96 well plate setup is shown in Figure 2.7 (modified from [Klatte2009]). The temperature profile of the qPCR reaction is shown in Table 2.6.

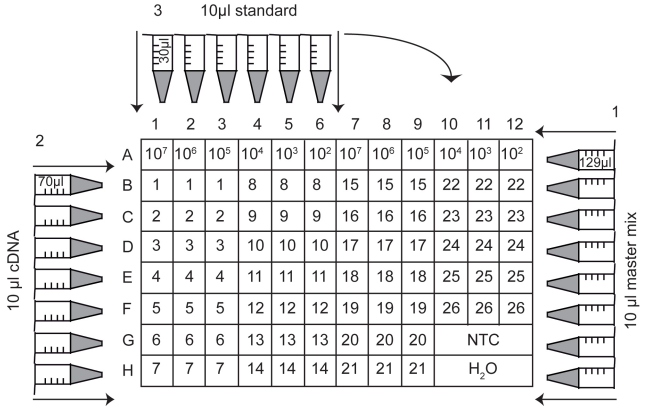


Figure 2.7: Pipetting scheme for qPCR analysis as described in Klatte and coworkers 2009 ([Klatte2009], modified). First, the master mix was aliquoted to 8 tubes (129 μl each tube) and transfered into the 96 well plate with the multichannel pipette. Second, the cDNA (B-H) and standards (A) were transferred into the rows. (NTC) negative (non-template) control from the cDNA synthesis. (H_2O) water control.

The 96 well plate was closed with a qPCR foil and centrifuged briefly just before PCR reaction.

Standard Preparation for qPCR DNA standards were used for the absolute quantification of copy numbers of DNA. In the first step, the DNA standards were generated. A PCR reaction was performed on template DNA (pMDC83-*SP-$_{PRP4}$-perlucin* K1) to amplify the *perlucin* gene. To achieve this, the primers 5-Prp4-Perlucin and 3-Perlucin were used in a PCR reaction prepared as described in Table 2.7 and the corresponding temperature profile is shown in Table 2.8. The PCR fragments were separated on a 1.5 % agarose gel (section 2.2.1.4). The DNA fragments were extracted with the Macherey and Nagel NucleoSpin^R ExtraktII Kit from the agarose gel. Finally, the DNA was eluted in a volume of 50 μl and the concentration was determined according to section 2.2.1.3. The copy number of DNA fragments was calculated using formula 2.1 and 2.2.

Table 2.7: Master mix for PCR reaction.

Master Mix	Amount [μl]
H_2O nuclease free add to	10
Phusion 2 x	10
5-Primer	1
3-Primer	1
DNA 500 ng	x

Table 2.8: Temperature profile for the standard preparation for the qPCR.

Step	Temperature [°C]	Time	Repeats
Denaturation	98	30 sec	
Denaturation	98	10 sec	
Annealing	60	30 sec	35
Elongation	72	15 sec	
Elongation	72	7 min	

$$\frac{copies}{\mu l} = DNA\ \frac{\mu g}{\mu l} \cdot \frac{1\ nmol \cdot bp}{660\ ng} \cdot \frac{1}{DNA\ bp} \cdot \frac{1\ mol}{10^9\ nmol} \cdot \frac{6.023 \cdot 10^{23}\ molecules}{mol} \quad (2.1)$$

$$\frac{copies}{\mu l} = DNA\ \frac{\mu g}{\mu l} \cdot \frac{1\ bp}{0.6\ \mu g} \cdot \frac{6.023 \cdot 10^{23}\ molecules}{DNA\ bp} \quad (2.2)$$

The DNA was adjusted to a copy number of 1.5 x 10^9 in x μl. A dilution series was prepared as described in Table 2.9. All standards were aliquoted to 30 μl in standard PCR strips and stored at -20 °C.

Table 2.9: Dilution series for qPCR standard preparation.

	DNA copy number [μl]	water nuclease free [μl]
(7) 10^7 copies / 10 μl	1.5 x 10^9 in x	1500 - x
(6) 10^6 copies / 10 μl	100 of (7)	900
(5) 10^5 copies / 10 μl	100 of (6)	900
(4) 10^4 copies / 10 μl	100 of (5)	900
(3) 10^3 copies / 10 μl	100 of (4)	900
(2) 10^2 copies / 10 μl	100 of (3)	900

Primer Matrix A primer matrix was performed to obtain the optimal primer concentrations for the qPCR reaction (Figure 2.8, modified after [Klatte2009]). The master mix for the PCR was composed as follows:

Reaction (master mix)	1 x
TaKaRa Premix 2 x	10 μl
Standard 10^4	0.2 μl
SYBR green	0.1 μl

The temperature profile for the reaction was identical to the profile of the qPCR reaction (Table 2.6). The Ct values were obtained from the reaction and plotted against the primer ratio. The lowest Ct value indicated the optimal primer concentration as shown in Figure 2.8 (modified from [Klatte2009]) (C) for the perlucin primer pair (5-RT-Perlucin, 3-RT-Perlucin) with the ratio of 2:1.

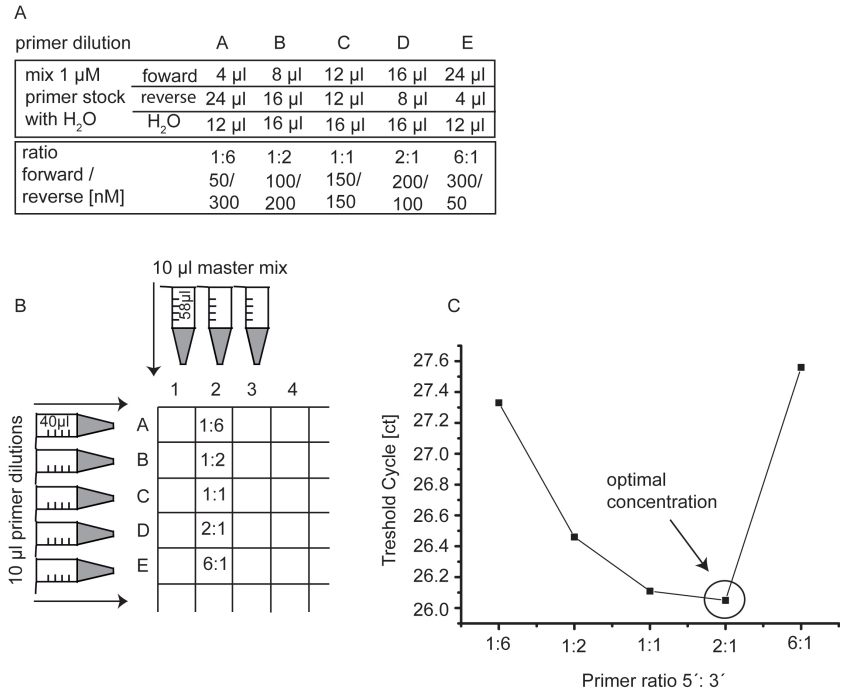


Figure 2.8: Experimental setup for primer optimization (modified from [Klatte2009]), (A) primer dilution series, (B) pipetting scheme, (C) data analysis: plot of the Ct values against the primer ratios. The lowest Ct values indicated the optimal primer concentration for the qPCR reaction as shown here for the perlucin primer (5-RT-Perlucin, 3-RT-Perlucin).

2.2.1.7 Restriction Digestion of DNA

The digestion of DNA was incubated at 37 °C for 1 h after combining the restriction enzyme with the corresponding buffer. The condition for the inactivation was according to manufactures guidelines (Fermentas, New England Biolabs). A standard reaction for a restriction digestion was accomplished as follows:

Reaction	1 x
DNA	0.5 μg
Buffer 10 x	1 μl
Restriction enzyme	2.5-5 units
H_2O nuclease free	add to 10 μl

2.2.1.8 Ligation

Digested DNA fragments and plasmid DNA were separated in an agarose gel and cleaned up with the NucleoSpin^R ExtraktII Kit from Macherey and Nagel. Isolated DNA fragments were incubated for ligation reaction with the ligase enzyme and the corresponding buffer at 16 °C over night. The ratio (amount of molecules) of a linear PCR fragment and the vector DNA was optimized to 5:1. The enzyme was inactivated by heat treatment after the ligation for 20 min at 65 °C. A standard reaction for a restriction digestion was accomplished as follows:

Reaction	1 x
Ligation buffer	1 μl
Vector DNA digested	x μl
PCR fragment digested	x μl
T4 ligase	1 μl
H_2O nuclease free	add to 10 μl

The ligation reaction was transformed into XL1-Blue bacteria cells (Stratagene).

2.2.1.9 Gateway System

All constructs for protein expression in plants were designed and adapted to the gateway system. The gateway system allows an efficient transfer of DNA strands from one plasmid to another. The gateway system is based on a recombination process from the

lambda phage that varies from standard cloning procedures. The system has a high positive transformation rate, up to 99 %. This is guaranteed through a suicide gene, which prevents cell growth without a recombination process. The system is divided in three steps. The first step involves a PCR reaction with specific primers in order to amplifiy the gene of interest and to add the recombination sites. In the second step, the PCR fragment will recombine with a so-called entry vector (BP reaction). In the third step, the DNA fragment of interest will be transferred via recombination to a selected destination vector (LR reaction). The gene synthesis and the subcloning into the entry vector pENTR/D-TOPO was performed by Entelechon. The LR reaction was performed as follows: 7 μl of the entry clone (50-150 ng/μl) was combined with 1 μl of the destination vector (150 ng/μl) in a clean Eppendorf tube. The clonase enzyme mix was shortly mixed and 2 μl of the mix was added to the reaction. The reaction was incubated over night at 25 °C and finished due to the addition of 1 μl proteinase K and an incubation step for 10 min at 37 °C. The final reaction mix was transformed into One Shot TOP10 cells (Invitrogen) for vector amplification. Gateway vectors without insert were reproduced in One Shot Survival T1-Phage resistance cells (Invitrogen).

2.2.2 Microbiological Methods

2.2.2.1 Production of Chemical Competent E. coli Cells

4 ml of LB medium was inoculated with a single colony of the bacteria cells and grown for 16 h at 37 °C and 180 rpm. 100 ml of LB medium was inoculated with the over night culture and allowed them to grow to a final O.D.600 of 1. The cells were centrifuged at 4 °C and 3,000xg for 10 min and subsequently kept on ice. The supernatant was discarded and the pellet was resuspended carefully with the pipette in ice-cold 0.1 M $CaCl_2$ followed by a 30 min incubation on ice. The cells were centrifuged as described above and resuspended in 4 ml of 15 % glycerol, 0.05 M $CaCl_2$. Finally, the cells were aliquoted on ice to 200 μl, shock frozen in liquid nitrogen and stored at -80 °C until use.

2.2.2.2 Transformation of Bacteria

Escherichia coli Chemical competent *E. coli* cells were incubated on ice for 15 min with 3 μl of vector DNA (approximately 1.5 μg) and subsequently transformed with a heatshock for 30 sec at 42 °C in a preheated water bath. The cells were immediately incubated for 2 min on ice before the addition of 1 ml of LB medium and a further

incubation at 37 °C for 30 min with 180 rpm. Transformed bacteria are finally plated on LB agar plates containing the corresponding antibiotics (concentrations are listed in Table 2.1), grown over night and selected for analysis.

Agrobacterium tumefaciens 3 ml of LB medium was inoculated with a single colony of *Agrobacterium tumefaciens* and grown for 16 h with rifampicin (final concentration 50 μg/ml) and carbenicillin (final concentration 34 μg/ml) in the LB medium at 28 °C and 220 rpm. 50 ml of LB was inoculated with 500 μl over night culture and allowed the cells to grow to a final O.D.600 of 0.8. The culture was incubated for 10 min on ice and centrifuged at 4 °C and 3,000xg to pelletize the cells and the remaining cell pellet was washed once with 10 ml of ice cold 20 mM $CaCl_2$. Cells were centrifuged again as described above and the pellet was resuspended in 1 ml of ice cold 20 mM $CaCl_2$ and aliquoted to 100 μl in sterile Eppendorf tubes. 1 μg of the vector DNA was added to each aliquot and subsequently frozen in liquid nitrogen. The tubes were incubated at 37 °C for 5 min before 1 ml of LB medium was added to each tube and shaked for 2 h at 220 rpm and 28 °C. Bacteria were centrifuged at 1,000xg for 1 min and the supernatant was removed to a final volume of 100 μl. The cells were resuspended and plated on inverted LB agar plates with the corresponding antibiotics (concentrations are listed in Table 2.1). The clones were grown inverted at 28 °C for 2 d and then selected for further analysis.

2.2.2.3 Growth of Bacteria for the Recombinant Protein Expression

5 ml of LB medium (final concentration of 100 μg/μl ampicillin) was inoculated with a single colony of the bacteria and grown over night at 37 °C and 180 rpm. 100 ml of LB medium with ampicillin was inoculated with one over night culture for the expression of GFP-perlucin and perlucin after 14-16 h. 250 ml of LB medium with ampicillin was inoculated with 2 single over night cultures for the expression of GFP. The induction of protein expression was achieved with the addition of IPTG (final concentration of 1 mM for standard conditions and 0.5 mM for optimized conditions) at an O.D.600 of 0.5-0.7. The growth conditions were optimized for each protein and will be described in the result part. 100 μl of each culture were harvested in 1 h intervals (t0-t4) and after 14-16 h (t5-t6), centrifuged at 4 °C, 1,200xg for 15 min for protein expression analysis. The pellet was resuspended in 5 μl water, mixed with equal amounts of sample buffer (2 x Lämmli) and stored at -20 °C until SDS-PAGE analysis.

2.2.2.4 Growth of Agrobacterium tumefaciens for Tobacco Infiltration

2 ml of LB medium with antibiotics (rifampicin, carbenicillin and the antibiotic for the plasmid with the gene of interest, Table 2.1) was inoculated with a single colony of *Agrobacteria* from a fresh agar plate. The bacteria were grown to an O.D.600 of 1.5 and harvested by centrifugation (5,000xg, 4 °C, 15 min). The pellet was washed once in infiltration medium (2 mM NaH_2PO_4, 50 mM MES buffer (pH 5.6), 0.5 % glucose). The bacteria were activated for 1 h at RT after the addition of acetosyringon to a final concentration of 100 μM before infiltration (section 2.2.4.6).

2.2.2.5 Growth of Agrobacterium tumefaciens for the Floral Dip

5 ml of LB medium with antibiotics (rifampicin, carbenicillin and the antibiotic for the plasmid with the gene of interest, Table 2.1) were inoculated with a single colony of *Agrobacteria* from a fresh agar plate. 100 ml of LB medium with antibiotics were inoculated with 2 ml of the over night culture (grown for 14-16 h). The bacteria were grown to an O.D.600 of 1.5, harvested by centrifugation (5,000xg, 4 °C, 15 min) and washed once in infiltration medium (10 mM $MgCl_2$, 10 mM MES/KOH (pH 5.2), 5 % sucrose, 0.05 % silwet gold). For the floral dip (section 2.2.4.7), the bacteria were activated for 3 h at RT after the addition of acetosyringon to a final concentration of 100 μM.

2.2.2.6 Preparation of Glycerol Stocks

Long term storage of bacteria cultures were achieved by glycerol stocks preparation. To accomplish this, 800 μl of an over night cell culture and 400 μl of 70 % sterile glycerol were briefly mixed in a cryo tube. The tube was immediately shock frozen in liquid nitrogen and stored at -80 °C.

2.2.3 Protein Methods

2.2.3.1 Protein Extraction from Plant Material

The plant material was gound to powder under liquid nitrogen. 100 μl of protein extraction buffer (750 mM Tris/Cl (pH 8.0), 4 % glycerol, 100 mM β-Mercaptoethanol, 0.1 % EDTA, 2 % SDS) was added to 100 μg of the powdered plant material in a reaction tube on ice followed by vigourous mixing and incubation on ice for 15 min. Subsequently, the

extract was centrifuged (4 °C, 13,000xg for 15 min). The supernatant was transferred to a fresh tube and the protein concentration was quantified with a Bradford assay (1 ml cuvette, section 2.2.3.4). The supernatant as well as the pellet fraction were analyzed on a SDS-PAGE and Western blot for protein detection.

2.2.3.2 Extraction and Batch Purification of Recombinant Expressed Protein

Bacteria for recombinant protein extraction were grown as described in section 2.2.2.3. The cells were centrifuged in 50 ml falcon tubes at 4,000xg for 10 min at 8 °C. For the purification of GFP-perlucin, it was necessary to freeze the pellets once in liquid nitrogen before cell lysis. The supernatant was discarded and the pellets were resuspendend in lysis buffer (2 ml per 0.5 g of pellet weight). Lysozyme was added to a final concentration of 1 mM and the cell pellet was incubated for 30 min on ice. The cell lysate was centrifuged at 16,000xg for 45 min at 4 °C and the supernatant (crude extract) was sterile filtrated (0.45 μm and 0.22 μm) and used for purification.

Native protein crude extract from 50 ml *E. coli* culture was incubated with 50 μl Ni-NTA matrix for 1 h at 4 °C in a 50 ml falcon under rotation. The Ni-NTA matrix was washed twice with H_2O_{MP} and once with lysis buffer prior to use. The agarose beads were transferred to 1 ml to 2.5 ml Mobicol columns (with a filter paper of 10 μm pore size). The beads were washed with 5 column volumes (CV) of protein washing buffer to remove unbounded or weakly bounded unspecific protein. This step was performed in a Table centrifuge (1 ml column) or by air flow (2.5 ml column). The recombinant protein was eluted with at least 2 CV of elution buffer. The protein was analyzed on a SDS-PAGE and stored at 4 °C for short term storage. For long time storage, the protein solution was shock frozen in liquid nitrogen and kept at -80 °C . The composition of all solutions used for the purification of recombinant protein are listed in section 2.1.13.

2.2.3.3 Affinity Chromatography of Recombinant Expressed Protein

Recombinant proteins were purified with an Äkta Purifier system equipped with a His-trap affinity purification column of 1 ml volume (GE Healtcare). The run profile is described in Table 2.10 and the purified protein solution was analyzed on a SDS-PAGE prior to use and stored at 4 °C.

Table 2.10: Run profile of affinity purification chromatography of recombinant proteins with the Äkta Purifier system equipped with a 1 ml His-Trap column. (CV) Column volumes, (A) lysis buffer, (B) elution buffer.

Parameter	Settings
Detection wavelength	280 nm
Flow rate	1 ml/min
Equilibration	5 CV (100 % A, 0 % B)
Washing	5 CV (98 % A, 2 %B)
Gradient	in 5 CV (90 % A, 10 % B)
Elution	5 CV (0 % A, 100 % B)
Fraction size	1 ml
Reequilibration of the column	5 CV (100 % A, 0 % B)

2.2.3.4 Protein Quantification with Bradford Reagent

Standard quantification of recombinant protein was performed in a 96 well plate. 5 μl of BSA standard proteins (0, 0.1, 0.25, 0.5, 0.75, 1, 1.4 mg/ml) and the unknown samples was mixed with 250 μl of Bradford reagent (Sigma Aldrich), each in a separate well. The plate was incubated for 5 min at RT. The extinction was measured at a wavelength of 595 nm in a plate reader (Molecular Devices, Spectra Max 190). If the protein concentration was higher than 1.4 mg/ml, the samples had to be diluted for accurate measurements. Plant protein extracts were quantified in a 1 ml cuvette to avoid precipitation of Bradford chemical reagent with plant substances. Therefore, 1 μl of the protein samples (standard and unknown) were mixed with 799 μl of deionized H_2O and 200 μl of Bradford reagent. The samples were analyzed in a standard photometer (Eppendorf) as described above.

2.2.3.5 Dialysis and Concentration of Protein Samples

Dialysis tubes (Roth) were used with a 4-6 kDa molecular weight cut off (MWCO) for volumes exceeding 500 μl. The dialysis tubes were incubated for 20 min in H_2O_{MP} before adding the protein solution. The incubation was performed in 1 up to 2 liter solution at 4 °C for 24 h. One buffer change was performed. Volumes smaller than 500 μl were dialyzed in Mini Dialyzer units with a MWCO of 7-10 kDa (Fisher). Microsep

advanced centrifugal devices (Pall) were used for protein concentration. The protein samples were centrifuged for 45 min with 4,000xg (this step was repeated according to requirements). The protein solution was transferred to an Eppendorf tube and stored at 4 °C for short term storage. The protein solutions were shock frozen in liquid nitrogen and kept at -80 °C for long time storage.

2.2.3.6 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The SDS-PAGE allows a size separation of proteins in an electric field. All proteins were first negatively charged due to binding of SDS to the proteins. The proteins were focused in a stacking gel buffered to pH 6.8, then separated in a separation gel buffered to pH 8.8 (see for gel preparation Table 2.4). Protein extracts from plant material were loaded in an amount of 25 μg per lane (maximum). Dependent on the protein concentration of each sample, the proteins were mixed with 2 x or 4 x Lämmli buffer. All samples were heated prior to use for 5 min at 95 °C. 10 μl of the protein molecular weight standard was loaded on each gel (Spectra BR ^{TM}, Figure 2.6, left). Standard running conditions were 20 min at 80 V and for 85 min at 135 V. The gels were coomassie stained or used for the Western blotting procedure.

2.2.3.7 Silver Staining of Protein Gels

The silver staining was performed according to Blum et al. [Blum1987]. The gel was incubated for 1 h in fixation solution after standard SDS-PAGE analysis. The gel was washed twice (10 min in 50 % ethanol and 10 min in 30 % ethanol). Subsequently, the gel was incubated for exactly 1 min in sodium thiosulphate solution and washed three times for 20 sec in H_2O_{MP}. The gel was laid for 20 min in impregnation solution and washed again as described above. The protein bands were stained in developer solution until the background changed colour and the reaction was aborted with the stopper solution. The composition of each solution is listed in section 2.1.13.

2.2.3.8 Native Page Analysis

10 % TGX gels (Bio-Rad) were used for native protein analysis. The protein samples were mixed at a ratio of 1:1 with native sample buffer and the run conditions were 30 min at 200 V in native running buffer (the buffer preparation is described in section 2.1.13). The proteins were extracted from the gel through diffusion in equal amounts

of 10 mM Tris/Cl, pH 8.7. The eluate was used for small volume calcium carbonate precipitation assays (described in section 2.2.7.2).

2.2.3.9 Western Blotting

Proteins were transferred from the SDS gel to a PVDF membrane (Millipore) at 100 V for 1 h in transfer buffer. Two different types of developmental procedure were used (enhanced chemiluminescence (ECL) and alkaline phosphatase (AP)) after blocking of unspecific binding sites in 5 % milk powder over night at 4 °C. The blot was washed in 1 x TBST buffer (3 x 15 sec and 2 x 10 min) and the membrane was incubated in the first antibody for 2 h at room temperature (RT). For the first antibody, polyclonal anti-GFP from rabbit or anti HIS_6 (Dianova, 1:2,000), was used for alkaline phosphatase (AP) development and a monoclonal anti-GFP from mouse (Roche, 1:20,000) for ECL development.

Subsequently, the blot was washed in 1 x TBST buffer (3 x 15 sec and 2 x 10 min) before the secondary antibody was added. Secondary antibody incubation was performed for 1 h at RT with AP-conjugated anti rabbit antibody (Dianova, 1:5,000) and HRP-conjugated anti mouse antibody (horseradish peroxidase, Sigma 1:100,000) for AP and ECL detection. After secondary antibody incubation, the blot was washed in 1 x TBST buffer (3 x 15 sec and 2 x 10 min).

For ECL detection, the solutions from the Western blotting detection kit were mixed at a ratio of 1:1 (each 500 μl) and incubated with the blot. The detection of chemiluminescence (ECL, GE Healthcare) was recorded with a bioillumination system (Biozyme) over time periods of 5, 15 and 30 min.

For the AP-detection, the blot was incubated for an additional minute in TNK buffer and 10 min in AP buffer. The blot was developed for up to 30 min in 5 ml of AP buffer containing 2.5 μl of NBT and 15 μl of BrCIP. The reaction was stopped with H_2O_{MP}. In order to evaluate a successful transfer of the proteins to the membrane, blots were stained with Ponceau-rouge (Merck). The composition of each solution is listed in section 2.1.13.

2.2.3.10 Sample Preparation for Mass Spectrometric Protein Analysis

Standard SDS-PAGE was performed with coomassie staining. The gel was de-stained in 7 % acetic acid solution for 3 h and was incubated in water over night (the solution was changed occasionally). The bands of interest and a control piece without protein

were excised and cut into pieces of 2 mm^2 size with a sharp scalpel. The pieces were transferred to a fresh 2 ml Eppendorf tube.

Subsequently, they were extracted with 50 mM NH_4HCO_3, 50 mM NH_4CO_3/25 % acetonitrile, 25 % acetonitrile, 50 % acetonitrile (each 30 min under rotation at RT). The pieces were lyophilised for 1 h and then incubated in a solution of 2 μg trypsin/100 μl gel volume in 50 mM NH_4CO_3 at 37 °C over night. The protein fragments were extracted with 100 mM NH_4HCO_3 and mass analysis was performed by the group of Prof. Deutzmann (University of Regensburg) and Prof. Heinzle (K. Hollemeyer, University of Saarland).

2.2.3.11 Reisolation of Protein from Calcium Carbonate Precipitates

Calcium carbonate precipitates from large volume assays (section 2.2.7) were pooled and centrifuged at 3,200xg for 10 min at 4 °C. The pellet was transferred to a dialysis tube with a MWCO of 4-6 kDa. The mineral phase was dissolved in 10 % acetic acid and the remaining organic part was dialysed against 10 mM Tris pH 8.7. Samples were lyophilised and resuspended in 100 μl of 10 mM Tris pH 8.7 and analysed with SDS-PAGE (section 2.2.3.6).

2.2.3.12 Field-Flow Fractionation Analysis of Purified Recombinant Protein

The asymmetric field-flow fractionation system (AF4 (or FFF), Wyatt) was used to separate proteins in dependence on their size in a separation channel (Figure 2.9, top: Schematic overview, bottom: Sample separation channel). The system allows the investigation of the agglomeration behaviour of the recombinant proteins in different solutions. The AF4 system consists of a high pressure liquid chromatography pump (HPLC, Agilent) equipped with an autosampler, UV and fluorescence detector (FL). An Eclipse instrument (Wyatt) controls the fluidics of the system. A light scattering unit (Dawn Heleos II for static light scattering (LS)) and dynamic light scattering (DLS, QELS), Wyatt) can be used for detection and hydrodynamic diameter calculation of single protein species and their agglomerates after batch purification (batch purification is described in section 2.2.3.2). Schematic representation of the AF4 system is shown in Figure 2.9, top.

An overview about the fluidics in the separation channel is shown in Figure 2.9, bottom: The sample is injected into the channel (A) and focused in a specific area of the separation channel (B). The sample is eluted and separated after size due to the activated cross

flow such that the elution of larger agglomerates is retarded (C). All separated particle fractions are verified with the different detector sources.

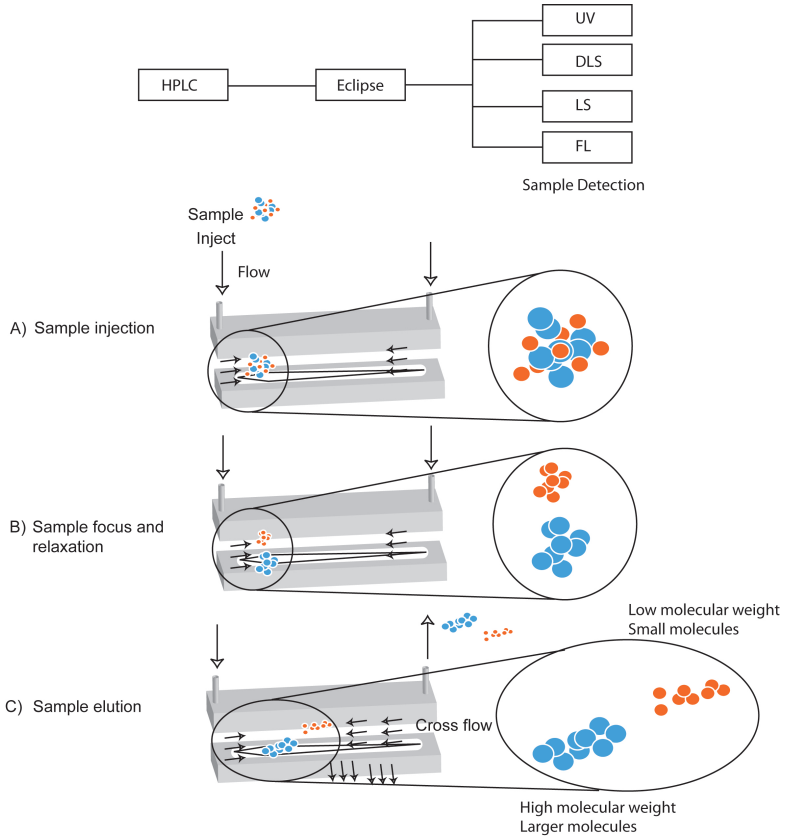


Figure 2.9: Schematic overview about the AF4 system. The HPLC compound and the Eclipse device controls the fluidics in the channel. (A) The sample is injected into the channel, (B) the sample is focused and relaxed such that the single particles are arranged in dependence on their size in the focus area, (C) sample separation and elution is under control of the cross flow and the different fractions were detected with the detector sources ((UV) detector, (DLS) dynamic light scattering unit, (LS) static light scattering unit and (FL) fluorescence detector).

The light scattering signals from the DLS and LS were recorded by the ASTRA software and the hydrodynamic radius (Rh) was automatically calculated from the diffusion coefficient (equation 2.3). The diffusion coefficient was determined by the autocorrelation function as described in the equation 2.4. The UV detector was used to determine the concentration of the sample and the FL detector allowed to record signals derived from the GFP protein. The system was established in cooperation with Gabi Klein and Dr.

Tobias Kraus (Structure Formation Group).

$$R_h = \frac{k_b T}{6\pi\eta D_t} \tag{2.3}$$

k_b	Boltzman constant
T	temperature
η	viscosity of solvent
D_t	diffusion coefficient
R_h	hydrodynamic radius

$$g^{(2)}(\tau) = 1 + \beta e^{-2D_t q^2 \tau} \tag{2.4}$$

τ	delay time
β	correlation function amplitude at zero delay
q^2	scattering vector

The protein monomers and dimers were with less than a 10 nm radius too small for a molar mass analysis. Therefore, only the hydrodynamic radius was determined. The run parameter for favourable separation processes were optimized for each recombinant protein and are specified in the appendix section 5.2.7. The protein samples were concentrated prior to use to obtain a significant QELS signal (GFP = 2 mg/ml and GFP-perlucin = 7 mg/ml, inject volume of 30 μl). All samples and buffers were 0.22 μm sterile filtrated and not used for longer than 3 days. The agglomeration behaviour was investigated in the precursor solutions for the inorganic (calcium carbonate) precipitation assay (section 2.1.13).

2.2.3.13 Size Exclusion Chromatography of Purified Recombinant Protein

The size exclusion chromatography (run parameter are listed in Table 2.11) was used to separate different protein species in the one step eluate after batch purification (described in section 2.2.3.2). All samples were injected in elution buffer, whereby lysis buffer was used as running buffer system. The purified protein fractions were analyzed on a SDS-PAGE prior to use. The composition of all solutions used for the chromatography are

registered in section 2.1.13. All solutions and buffers were sterile filtrated 0.22 μm and stored at 4 °C.

Table 2.11: Run profile of a size exclusion chromatography (Superose 12 10/300GL column). This chromatography method allows a separation of different protein species in the one step eluate after batch purification. (CV) Column volume.

Parameter	Settings
Detection wavelength	280 nm
Flow rate	0.4 ml/min
Equilibration	2 CV
Elution	1 CV
Fraction size	1 ml
Reequilibration	2 CV

2.2.3.14 Protein Preparation for Nano-Particle Synthesis

The protein preparation was accomplished according to section 2.2.3.2. The protein solution was concentrated in Pall centrifugal devices (3 kDa MWCO) to a final concentration of 15-25 mg/ml. The protein samples were supplied in L-arginine (1.29 g/l) or 10 mM $NaHCO_3$ solution, which was in dependence on the synthesis strategy. All solutions were sterile filtrated prior to use. The protein was washed 3x with 5 ml of the corresponding solution in the Pall centrifugal devices. The centrifugation steps were performed at 8 °C at 3,200xg for 45 min. The protein samples were stored at 4 °C.

2.2.4 Plant Biological Methods

2.2.4.1 Seed Sterilization with Sodium Hypochlorite

Seeds were incubated in 1 ml of seed sterilization solution (6 % sodium hypochlorite, 0.1 % triton-X-100) for 10 min and subsequently washed 5x in sterile H_2O_{MP}. In between each step, the seeds were sedimented with a short centrifugation step (10 sec, 5,000xg) and finally stored for 2 d at 4 °C for stratification or plated on agar plates for further analysis as described in the sections 2.2.4.3 and 2.2.4.8.

2.2.4.2 Growth of Nicotiana benthamiana and Arabidopsis thaliana on Soil

Nicotiana benthamiana was grown on a mixture of soil and vermiculite (75 % turf and 25 % vermiculite klassmann substrate 2). The *Arabidopsis thaliana* plants were grown on turf balls. All seeds were sterilized before seeding (section 2.2.4.1). The soil was baked at 80 °C for 24 h before use.

2.2.4.3 Growth of Arabidopsis thaliana on Agar Plates

All seeds were sterilized with sodium hypochlorite (section 2.2.4.1) and plated in rows (12-16 seeds/row) on MS agar (4.3 g/L MS salt, 1 % sucrose, plant agar 0.8 %) plates (Greiner). The plates were closed with parafilm and put into home-made racks. The plants were allowed to grow in the plant growth chambers under following conditions:

- 22 °C day temperature (16 h, 120 μm^2 s^{-1}, humidity of 70 %)
- 18 °C night temperature (8 h, 0 μm^2 s^{-1}, humidity of 70 %)

2.2.4.4 Root Growth and Germination Assay

Plant seeds were sterilized as described in section 2.2.4.1 and plated on MS agar in two rows with 12 seeds per row on each plate. The plants were grown for 14 days (16 h: 22 °C, day period with 150 μm^2 s^{-1} and 8 h: 18 °C 0 μm^2 s^{-1}, 70 % humidity). The germination rate was determined and the lengths of their roots were measured.

2.2.4.5 T-DNA Insertion

Transfer DNA (T-DNA) is a standard tool to transform plants with foreign DNA fragments and to generate trangenic lines. The T-DNA is a part of a vector DNA, which originates from the soil bacterium *Agrobacterium tumefaciens*. These bacteria infect naturally dicotyledons and use the T-DNA to transfer its own DNA to the plant cells, which is flanked by two sequence fragments. The palliative fragments are called left border (LB) and right border (RB) and are essentiel for the transformation process. A schematic overview in Figure 2.10 shows the transformation procedure, whereby virulence genes (vir genes) are necessary for the DNA transfer.

Wounded plants are able to produce phenolic compounds, which induces the synthesis of vir proteins (acetosyringon mimicks this compound in the laboratory setup). These vir proteins allows to process the T-DNA for transfer (t-strand, Figure 2.10). The

Agrobacterium transfers the t-strand and the vir proteins through channels to the plant cell and a t-complex is formed (Figure 2.10). Subsequently, the t-complex is tranfered to the nulcleus and the t-strand is inserted into the plant genome [Gelvin2005].

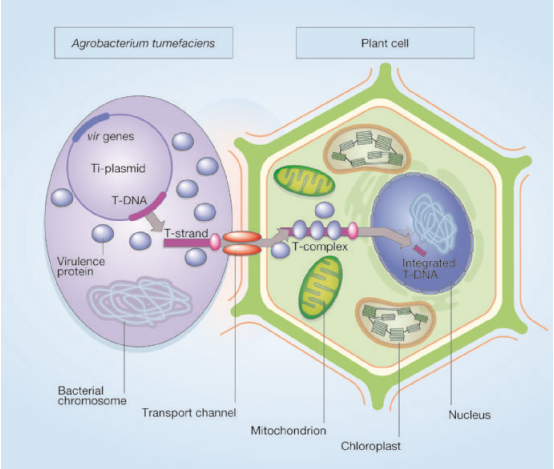


Figure 2.10: Schematic illustration of plant infection by the *Agrobacterium tumefaciens* [Gelvin2005]. The natural infection allows a genetic modification of plant cells with the DNA of interest. The gene of interest (as part of the t-DNA) will be transferred to the plant cell and inserted into the plant genome, which is enabled by the vir genes.

2.2.4.6 The Infiltration of Tobacco Leaves with Agrobacterium tumefaciens

The *Agrobacteria* cultures were grown as described in section 2.2.2.4 and the final bacteria solution was infiltrated with a blunt syringe into the tobacco leaves (3-4 leaf stage, bottom side). The protein expression was analyzed after 48-72 h.

2.2.4.7 The Floral Dip Method [Clough1998]

Closed inflorescences of approximately 40 plants were transformed with the *Agrobacterium tumefaciens* cells containing the construct of interest. How to growth the *Agrobacterium tumefaciens* cultures for the floral dip is described in 2.2.2.5 The inflorescences were dipped into the infiltration solution twice for 30 sec (10 mM $MgCl_2$, 10 mM MES/KOH (pH 5.2), 5 % Sucrose, 0.05 % Silwet gold). The infiltrated plants were covered with a foil for the next 24 h to increase the transformation rate. All siliques were harvested after their ripening and analyzed with regard to the inserted gene. Subsequently, the selection was performed on hygromycin containing MS medium and all

positive plants were genotyped (section 2.2.4.9) to verify the hygromycin selection (section 2.2.4.8).

2.2.4.8 Hygromycin Selection of Transformed Arabidopsis thaliana Plants [Harrison2006]

Sterilized seeds (section 2.2.4.1) were plated on hygromycin containing MS agar plates (25 μg/ml of hygromycin). The presence of the antibiotic allowed a selection of positive transformed plants. Transformed plants bore an increased length growth compared to untransformed seedlings and were distinguished by their elongated hypocotyl after one week. The plates were incubated under following conditions:

- 2 d, 0 μm^2 s^{-1} light at 4 °C
- 6 h, 120 μm^2 s^{-1} light at 22 °C
- 2 d, 0 μm^2 s^{-1} light at 22 °C
- 2 d, with day (16 h) and night cycle (8 h) at 22 °C

Selected seedlings were transferred to soil and allowed to grow further (section 2.2.4.2).

2.2.4.9 Genotyping of the Plant Material

Parental t0 seeds from fresh transformed plants were collected and plated after sterilization on hygromycin MS agar plates (growth conditions are described in section 2.2.4.8). The general selection procedure is shown in Figure 2.11. The DNA was extracted from each positive plant and analyzed by PCR, regarding the gene of interest. In the t1 generation, plants were grown with the genotype wild type/insert (wt/x) or wild type/wild type (wt/wt). Plants, which revealed a positive PCR signal with the correct molecular weight were selected (wt/x) for further cultivation and plants are discarded, if they had the genotype wt/wt. The t2 generation exhibited 3 different genotypes. Plants were homozygous for the insert (x/x), heterozygous for the insert (wt/x) or homozygous for the wild type (wt/wt). The analysis of the t2 progeny revealed the genotype of the mother plant. If all the progeny was positive in the genotyping PCR, the mother plant was homozygous (x/x) and can be used as working plant. The Extract-N-Amp^{TM} Plant Kit (Sigma Aldrich) was used for DNA extraction and two additionally performed PCR reactions were used to amplify the gene of interest. The corresponding temperature profiles for the PCR reaction is given in Table 2.12 (PCR1) and 2.13 (PCR2). Different primer were used for the genotyping procedure to evaluate the best primer combination as shown in Figure 2.12.

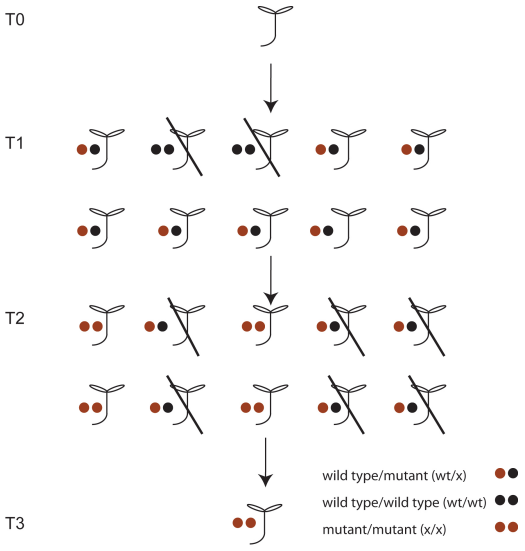


Figure 2.11: Selection procedure of transgenic *Arabidopsis thaliana* plants. The closed inflorescences of *Arabidopsis thaliana* were infiltrated with *Agrobacterium tumefaciens* cells containing the gene of interest and all seeds were selected (t0) and plated on hygromycin MS agar. Positive plants (wt/x) were transferred to soil. The presence of the insert was verified by PCR after DNA extraction. In the t1 generation, plants had the genotype wild type/insert (wt/x) or wild type/wild type (wt/wt). All positive plants were selected and allowed to grow further. The analysis of the t2 progeny exhibited which plant was homozygous for the insert (x/x). If all the progeny was positive in the PCR, the mother plant was homozygous (x/x) and can be used as working plant.

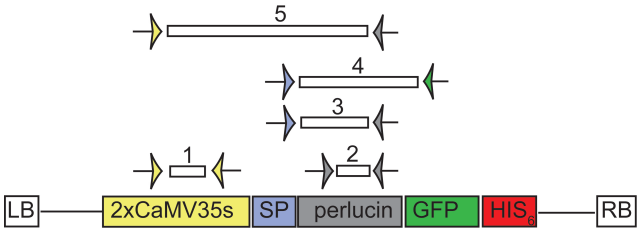


Figure 2.12: To verify positive transformed *Arabidopsis thaliana* plants, DNA was extracted from each plant and analysed by PCR. Different primer pairs were used to evaluate the best primer combination. Primers were selected that amplifiy the following regions: (1) within the 2xCaMV35s promotor region, (2) within the *perlucin* gene, (3) from the signal peptide to the *perlucin* gene, (4) from the signal peptide to the *GFP* gene, (5) from the 2xCaMV35s promotor region to the *perlucin* gene.

Table 2.12: Temperature profile for a standard touch-down PCR reaction after DNA extraction with the Extract-N-Amp^{TM} Plant Kit (Sigma Aldrich). The polymerase was directly delivered within the kit and the master mix was prepared according to manufactures guidelines.

Step	Temperature [°C]	Time	Repeats
Denaturation	94	3 min	
Denaturation	94	1 min	
Annealing	68 (-1)	1 min	10
Elongation	72	45 sec	
Denaturation	94	1 min	
Annealing	62	1 min	25
Elongation	72	45 sec	
Elongation	72	10 min	

Table 2.13: Temperature profile for the second PCR reaction after DNA extraction with the Extract-N-Amp^{TM} Plant Kit (Sigma Aldrich). The PCR reaction was performed to amplify the signals from the first reaction. A Dream Taq polymerase (Fermentas) was used for master mix preparation according to manufactures guidelines.

Step	Temperature [°C]	Time	Repeats
Denaturation	94	3 min	
Denaturation	94	1 min	
Annealing	62	1 min	25
Elongation	72	45 sec	
Elongation	72	10 min	

2.2.4.10 Plasmolysis

The plasmolysis experiment causes retraction of the protoplast from the cell wall and the cell vacuole through an osmotic lost of water. The plasmolysis was induced in tobacco

leaf cells due to an incubation in 1 M potassium nitrate (KNO_3) for 5-10 min. The reversible reduction of the protoplast size allowed a more detailed analysis of fluorescence pattern in the plant cell after transformation with the gene of interest. Fluorescence signals from cell wall and the inside of the cell were distinguished after the plasmolysis. Furthermore, cytoplasmic connections ("Hechtian stands"), which were anchored at the cell wall, were visible after plasmolysis, as shown in Figure 2.13.

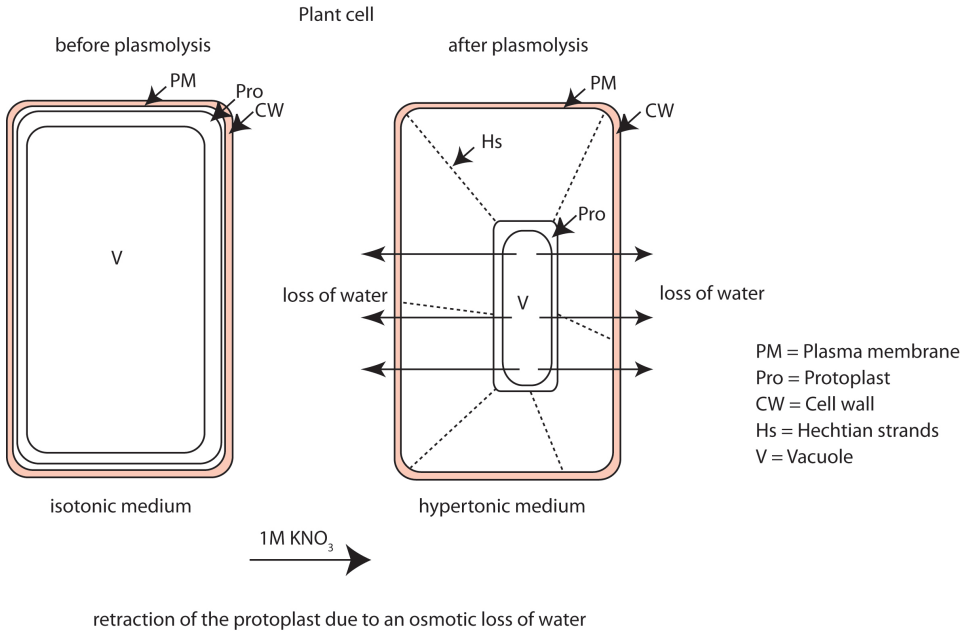


Figure 2.13: Schematic overview of a plasmolysis event in a plant cell induced due to an incubation in 1 M KNO_3. Cytoplasmic connections "Hechtian strands" are visible if the cell is plasmolysed. (PM) Plasma membrane, (P) protoplast, (CW) cell wall, (Hs) Hechtian stands, (V) vacuole.

2.2.5 Spectroscopic Methods

2.2.5.1 Raman Spectroscopy

Calcium carbonate polymorphs were characterized using a LabRam ARAMIS Raman microscope (Horiba Jobin Yvon, New Jersey, USA) equipped with the LabSpec Software package. Specimens from bulk assays were allowed to dry in air on stainless steel for imaging with the MPlan 50x/0.75 (Olympus, Hamburg, Germany) air objective. In order to reduce the possibility of recrystallization artifacts, precipitates were spectroscopically analyzed using a 100x/0.80 (Olympus, Hamburg, Germany) water immersion objective measured directly in solution on silicon wafer or in microscopic chambers with 785

nm Nd-YAG Laser, 1800 lattice over period of 20 sec. In order to match the contact area of the water objective, it was a required to fill the chambers of small volume assays with additional 10 μl fluid directly prior to taking the spectra. Spectra were automatically baseline adapted and background corrected with the LabSpec Software. Raman measurements were performed in collaboration with Christina Guth.

2.2.6 Microscopic Methods

2.2.6.1 Light Microscopy

LC-Polscope Microscopy All samples were prepared in water on glass slides, if not stated otherwise. An inverted light microscope Zeiss Observer Z1 equipped with an A Plan 10x/ 0.25 Ph1, LD Plan Neofluar 20x/ 0.4Korr Ph2 and a LD Plan Neofluar 344, 40x/ 0.6Korr Ph2 objectives and LC-PolScope technology (CRI Abrio Imaging System, L.O.T. Oriel, Darmstadt, Germany), including the respective filter sets was used to analyze precipitates with respect to birefringent retardance and orientation.

Fluorescence Microscopy Fluorescence images were taken using the same Zeiss Observer Z1 equipped with an A Plan 10x/0.25 Ph1, LD Plan Neofluar 20x/0.4Korr Ph2 and a LD Plan Neofluar 344, 40x/0.6Korr Ph2 objectives and with an epifluorescence Colibri illumination system combined with filter set 38 HE (Zeiss, Göttingen, Germany). Images were recorded with an AxioCamMRm (60N-C2/3 0.63x) digital camera and analyzed using the Zeiss AxioVision software (Zeiss, Göttingen, Germany). The exposure time was automatically adjusted using pre-defined algorithms.

Stereo Microscopy Morphological analysis of plant material was performed with a Leica M165C microscope, equipped with a DFC290 camera (Leica, Wetzlar, Germany).

Confocal Laser Scanning Microscopy (CLSM) The distribution of fluorescence signals within precipitates was investigated using a TSC SP2 confocal laser scanning microscope (Leica, Wetzlar, Germany) equipped with a HCX PL APO 40x/ 1.25-0.75 oil objective. Samples were mounted wet on glass slides. A 488 nm argon laser was used for GFP excitation and emission signals were recorded between 507-512 nm.

2.2.6.2 Electron Microscopy

Scanning Electron Microscopy An environmental scanning electron microscope (E-SEM Quanta 400, FEI, Eindhoven, the Netherlands) was used for structural surface characterization of precipitates. Sample droplets were directly mounted on a silicon wafer without additional coatings. Images were obtained in the low vacuum mode at 100 Pa pressure at 10-20 kV. SEM images were performed in collaboration with Magdalena Eder.

2.2.7 Calcium Carbonate Precipitation Assays

Two different calcium carbonate precipitation assays were developed based on previously established procedures with major modifications [Wheeler1981]. Protein samples (GFP, GFP-perlucin derivatives), and control proteins FITC-Concanavalin A were dialyzed against 10 mM Tris buffer pH 8.7 immediately prior to each assay to avoid early salt induced precipitation by protein purification buffers. All solutions were prepared with deionized water and sterile filtrated with a 0.22 μm PVDF filter. All glass equipment was thoroughly rinsed with deionised water immediately before use. All precipitation assays were accomplished at RT 22 °C +/- 2 °C.

2.2.7.1 Large Volume Assay

A volume of 6 ml 20 mM $CaCl_2$ solution (3 mM Tris pH 8.7) was supplemented without or with protein solutions of (a) GFP-perlucin: 5 μg, 25 μg, or 100 μg, b) GFP: 100 μg; c) Concanavalin A: 25 μg or 100 μg), 0.22μm filtrated and poured into 6 ml 20 mM $NaHCO_3$ (3 mM Tris pH 8.7) at constant rate on a Teflon-coated magnetic stirrer. The final protein concentrations in 12 ml were 0.4 μg/ml (5 μg), 2 μg/ml (25 μg) and 8 μg/ml (100 μg)

Assay I-$CaCl_2$ A volume of 6 ml 20 mM $CaCl_2$ solution (3 mM Tris pH 8.7) was supplemented without or with protein solutions of (a) GFP-perlucin: 5 μg, 25 μg, or 100 μg, b) GFP: 100μg; c) Concanavalin A: 25 μg or 100 μg), 0.22 μm filtrated and poured into 6 ml 20 mM $NaHCO_3$ (3 mM Tris pH 8.7) at constant rate on a Teflon-coated magnetic stirrer. The final protein concentrations in 12 ml were 0.4 μg/ml (5 μg), 2 μg/ml (25 μg) and 8 μg/ml (100 μg).

Assay II-$NaHCO_3$ A volume of 6 ml 20 mM $NaHCO_3$ solution (3 mM Tris pH 8.7) was supplemented without or with protein solutions of (a) GFP-perlucin 5 μg, 25 μg, or 100 μg, b) GFP: 100 μg; c) Concanavalin A: 25 μg or 100 μg), the solution was 0.22 μm filtrated and poured into 6 ml 20 mM $CaCl_2$ (3 mM Tris pH 8.7) at constant rate on a Teflon-coated magnetic stirrer. The final protein concentrations in 12 ml were 0.4 μg/ml (5 μg), 2 μg/ml (25 μg) and 8 μg/ml (100 μg). The pH values were recorded every 10 sec for at least 10 min from combining the solutions. A Prolab 3000 pH meter (SI Analytics) equipped with a N6000 electrode (SI Analytics) and MultiLab Pilot V5.04 software (SI Analytics) was used for automated recording of pH values.

2.2.7.2 Small Volume Assay

Small volume precipitation assays were performed in 30 μl volumes using disposable microscopy slides (Ibidi). A volume of 14.5 μl 20 mM $CaCl_2$ solution (3 mM Tris pH 8.7) was directly mixed in the chamber with 1 μl protein solution (1 μg/μl, 0.1 μg/μl and 0.01 μg/μl of GFP, GFP-perlucin derivatives). In the case of the full length GFP-perlucin, volumes of 1 μl were used directly after extraction from native polyacrylamide gels.

2.2.7.3 Statistical Analyis

Statistical analysis was performed with the students t-test.

3 Results

The result section consists of two major parts. The first part reports on the expression of the biomineralization proteins in a bacterial system and the second part refers to their expression in plant organisms.

3.1 Expression of Biomineralization Proteins in Bacteria

To achieve the biotechnological synthesis of biomineralization proteins (BP), *perlucin* was cloned for the heterologous expression in the bacterial system. Perlucin from the nacre (*Haliotis laevigata*) seemed to be a promising candidate for the synthesis of artificial biominerals due to its promoting effect on calcium carbonate precipitation, which was previously demonstrated *in vitro* (introduction 1.1.4.1). The final gene sequence was obtained by reverse translation from the protein amino acid sequence (UniProtKB accession no. P82596) and the codon usage was optimized using the Leto software package (Entelechon) to remove potential cleavage sites and to facilitate protein expression in the plant system, which will be described in the second part of the results.

3.1.1 Cloning Strategy

All codon optimized sequences used for transformation are listed in the appendix section 5.1. Two different strategies were chosen to achieve the expression of *perlucin* in bacteria. The *perlucin* gene was cloned in the first strategy (Figure 3.1, perlucin, construct A), whereas a *perlucin* gene with a *GFP* gene sequence was synthesized within the second strategy to obtain a *GFP-perlucin* sequence (Figure 3.1 construct B).

The GFP-fused proteins allowed protein tracking during its interaction with the mineral phase. However, it was also required to study the influence of the GFP protein on mineral formation. Both proteins contained a HIS_6 tag, for one step affinity purification of the proteins using a Ni-NTA resin.

In the first step, the *perlucin* gene was amplified by PCR, and *Kpn*I and *Pst*I restriciton sites were attached in the PCR reaction for subsequent cloning into the bacterial expression vector. The pENTR/-D-TOPO-SP_{EXT3}-*perlucin* vector served as template DNA. Different *Kpn*I 5-primers were designed due to the usage of two different expression vectors (construct A, pQE30 for perlucin expression and construct B, pQE31-*GFP-CBP* [Weiss2006] for GFP-perlucin, here the *CBP* was exchanged against the *perlucin* gene sequence, Figure 3.1). Both pQE30 and pQE31 sequences differed in one base pair and subsequently, the 5-Perlucin-KpnI primer was used for the synthesis of construct A and the 5-Perlucin-KpnI-pQE31 for construct B (the 3-Perlucin-PstI primer was identical for both constructs). A representative PCR result of the amplified *perlucin* gene of construct A is shown as example in Figure 3.2 A with a product length of 487 bp.

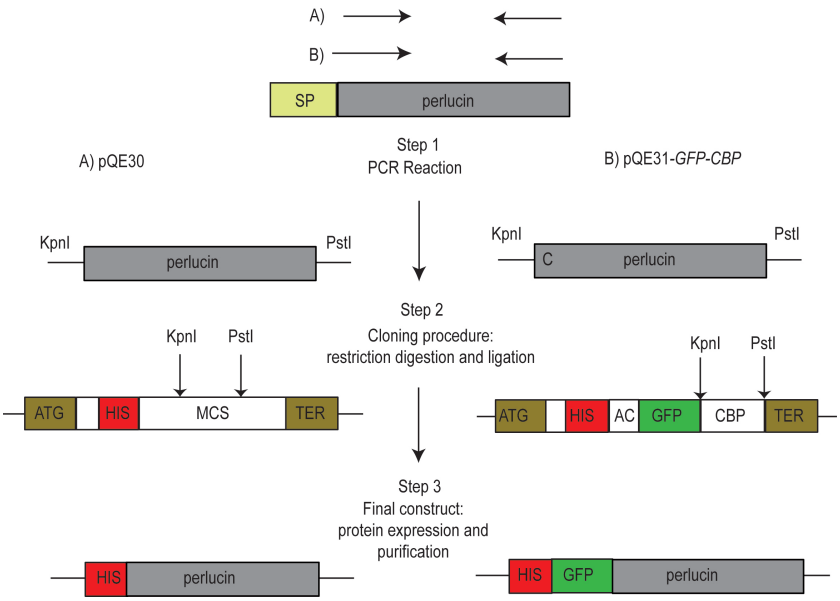


Figure 3.1: Cloning strategy for recombinant expression of perlucin (grey) with a HIS_6 tag (red) for one step affinity purification (pQE system from Qiagen). Strategy A included the production of the pure perlucin and strategy B the synthesis of perlucin with a N-terminal GFP fluorescence tag (green). The map of the expression vector is shown in Figure 2.3. (SP) Signal peptide, (ATG) start codon for transcription, (TER) terminator sequence for transcription, (MCS) multiple cloning site, (CBP) chitin binding protein.

The second step was to digest the expression vectors pQE30 and pQE31-*GFP-CBP* and the PCR products with *Kpn*I and *Pst*I (Figure 3.1, step 2). The vectors and the PCR products were purified via agarose gel and subsequently ligated to obtain the pQE30-*perlucin* and pQE31-*GFP-perlucin* expression vector. In construct B, the *CBP* gene

was exchanged against the *perlucin* gene (Figure 3.1 B, step 2). A restriction digestion was performed to control the presence of the insert. The obtained fragments correlated with the expected sizes. A double digestion with *Kpn*I or *Pst*I revealed a 3446 bp DNA fragment corresponding to the vector backbone and a 474 bp fragment representing the molecular size of the *perlucin* gene (Figure 3.2, B lane 1). A linearized pQE30-*perlucin* plasmid revealed a 3920 bp fragment after single digestion with *Kpn*I or *Pst*I (Figure 3.2, B lane 2-3) compared to the non-digested control (Figure 3.2, B lane 4). The final sequences in the vectors were verified by sequencing (5-pQE31-71 and 3-pQE31-195 primers). Construct A and B were transformed into XL1 *E. coli* cells for vector replication and protein expression. All applied methods are described in the material and method section 2.2.1.

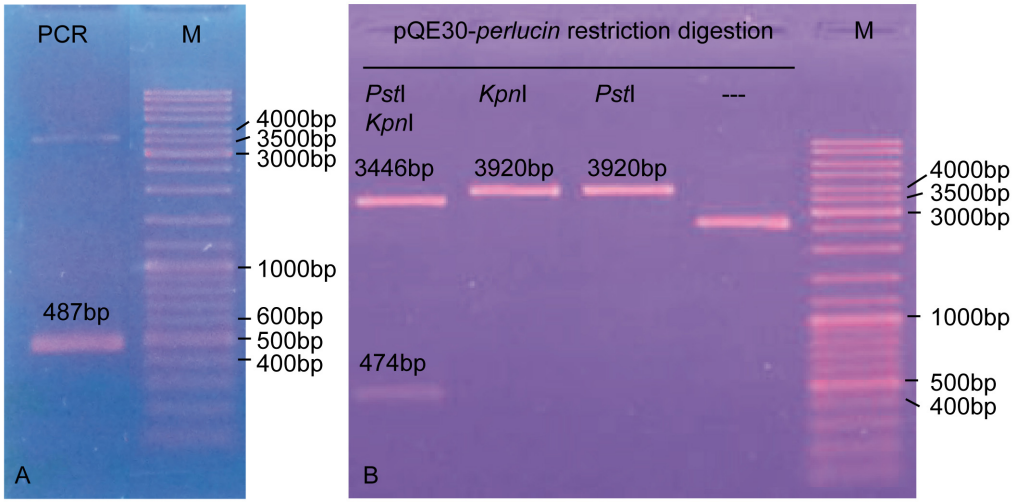


Figure 3.2: (A) Agarose gel showing PCR amplification of the *perlucin* gene after attaching *Kpn*I and *Pst*I restriction sites in the reaction, which resulted in a 487 bp DNA fragment. (B) The presence of two fragments after the double digestion of the final pQE-*perlucin* plasmid with *Kpn*I and *Pst*I (B, lane 1) revealed the successful ligation of the *perlucin* gene into the pQE30 vector. (—) Undigested pQE-*perlucin*, (M) molecular weight standard, GeneRuler^{TM}.

3.1.2 Protein Expression and Purification

3.1.2.1 Optimization of Protein Expression

After the cloning of the *perlucin* gene into bacterial expression system, the goal was to purify the recombinant protein under native conditions. Three different proteins were compared in this study: Perlucin (20.2 kDa, pQE30-*perlucin*), GFP-perlucin (46.8 kDa,

pQE31-*GFP-perlucin*) and the pure GFP (28.9 kDa, pGFP-*Sca-CDB*-8) as control. All proteins were purified with affinity Ni-NTA batch purification. The Äkta Purifier system was not used due to a continuous protein precipitation after the purification process (appendix Figure 5.7). The success of the purification was analysed with SDS-PAGE. GFP was very soluble (high protein yield), GFP-perlucin was slightly soluble (minor protein yield) and perlucin was insoluble under native extraction conditions (no detectable protein yield on SDS-PAGE, appendix Figure 5.2). To overcome the insolubility of perlucin and to enhance the expression of GFP-perlucin, the growth parameters were optimized. The standard conditions accomplished a growth at 37 °C over 5-6 hours and the protein expression was induced with 1 mM IPTG. Under optimized conditions, the temperature was reduced to 22 °C for 8 h and to 18 °C for 16 h. The IPTG concentration was narrowed to 0.5 mM for induction of protein expression (Figure 3.3 A shows the temperature profile for standard and optimized conditions).

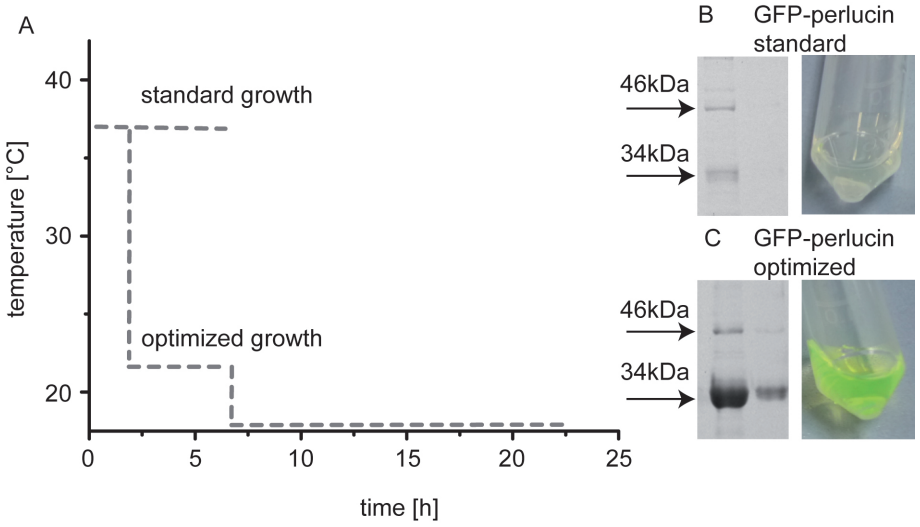


Figure 3.3: (A) Temperature profile for standard and optimized growth condition for bacteria cultures expressing GFP-perlucin. (B) The native protein yield from cultures grown under standard conditions was low. (C) Optimized culture conditions yielded about 10 x more soluble GFP-perlucin (46.8 kDa) and about 100 x more GFP-perlucin (34 kDa) as quantified by coomassie gels.

Furthermore, a second band for the GFP-perlucin protein detectable on the SDS-PAGE on a size of 34 kDa in addition to the expected size of 46.8 kDa (before and after optimization, Figure 3.3 B, C). The enhanced yield of recombinant protein after the

optimization procedure was visible on the stronger bands on the SDS-PAGE (mainly for the 34 kDa fragment) and the more fluorescent protein solution (Figure 3.3 B, C).

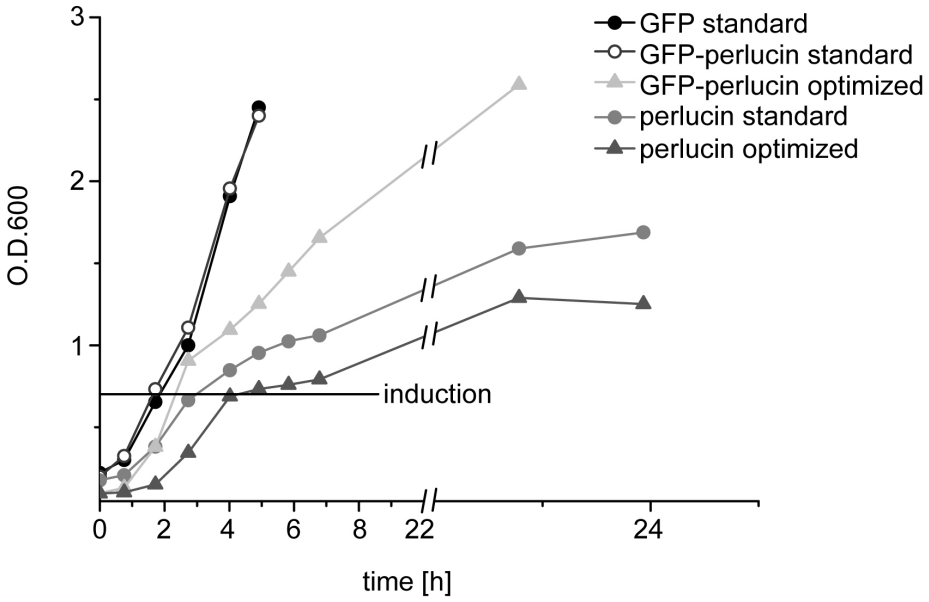


Figure 3.4: Cell density profiles for recombinant GFP, GFP-perlucin and perlucin expressing *E. coli* cultures. Standard culture conditions (circles) were suitable for GFP (black line, closed circles). GFP-perlucin containing cells reached a cell density (O.D.600) of ∼2.6 after 5 h (standard) and 24 h (optimized conditions, light grey triangles). Perlucin expressing cells divided slower under both, standard (grey circles) and optimized (dark grey triangles) growth conditions to a maximum O.D. of ∼1.6 after 24 h.

The expression of perlucin without GFP in *E.coli* led to a reduced growth rate reaching a final OD 600 of ∼1.8 (Figure 3.4, standard (grey circles) and optimized (dark grey triangles)), which was much lower when compared to *E.coli* strains expressing GFP and GFP-perlucin with a typical O.D.600 of ∼ 2.6 after 5 h (Figure 3.4, black line, open circles). SDS-PAGE analysis of perlucin revealed that the purification was not achieved either under standard or optimized conditions (appendix Figure 5.2, bottom row).

The corresponding images of purification steps from each construct are shown in the appendix Figure 5.1. A comparative set of SDS-PAGE analysis of GFP (standard), GFP-perlucin and perlucin (standard and optimized conditions) are shown in the appendix Figure 5.2 with a purity of $< 98\%$ for GFP and $< 95\%$ for GFP-perlucin as

estimated by the SDS-PAGE analysis. The growth of cell cultures and the purification of the recombinant proteins is described in section 2.2.2.3 and 2.2.3.2. In addition, the presence of the recombinant GFP-perlucin protein was also verified by Western blot using antibodies specific for GFP and the HIS_6 tag in the recombinant protein (appendix Figure 5.3). In order to produce native protein extracts for the biotechnological synthesis of artificial biominerals, the native insoluble perlucin protein was not analyzed further and only the native purified GFP and GFP-perlucin proteins were characterized.

3.1.3 Protein Characterization

3.1.3.1 Mass Spectrometric Analysis of Recombinant Proteins

Different protein species from the SDS-PAGE analysis were identified either with intact mass analysis (one step eluate) or after trypsin digestion (34 kDa and 46 kDa bands, Figure 3.3 B,C) by MALDI-TOF/TOF (MS+MS/MS). Intact mass analysis of GFP Ni-NTA eluates (Figure 3.5 A) revealed the expected molecular mass of 28.9 kDa compared to GFP-perlucin (Figure 3.5 B) that yielded in two major peaks at 29.8 kDa and 30.3 kDa, corresponding to truncated protein variants with 265 and 269 amino acids. A minor broad peak at 45 kDa indicated the presence of a protein variant with 395 amino acids. Protein dimer formation was indicated due to the presence of additional peaks in the intact mass spectra of GFP (57.9 kDa, Figure 3.5, A) and GFP-perlucin (59.4 kDa, Figure 3.5, B).

GFP-perlucin bands were digested with trypsin after SDS-PAGE separation (compare 46 kDa and 34 kDa band on the SDS-PAGE, 3.3 B,C and Figure 3.6). The presence of the two C-terminal repeat motives (SLHANLQQR) revealed the existence of the GFP-perlucin full length variant. Figure 3.6 A-E summarized the obtained results form MALDI-TOF/TOF (MS+MS/MS) analysis with a schematic illustration of all protein variants. Figure 3.6 represents the GFP protein (A), the truncated GFP-perlucins lacking 145 or 141 (B-C) amino acids and the full length variant including the two C-terminal repeat motives (E). The pure perlucin without GFP was not analyzed here due to its insolubility under native conditions (Figure 3.6 F). After identification of the different GFP-perlucin protein variants, the next objective was to separate them by native-PAGE and size exclusion chromatography and to study their native agglomeration behaviour.

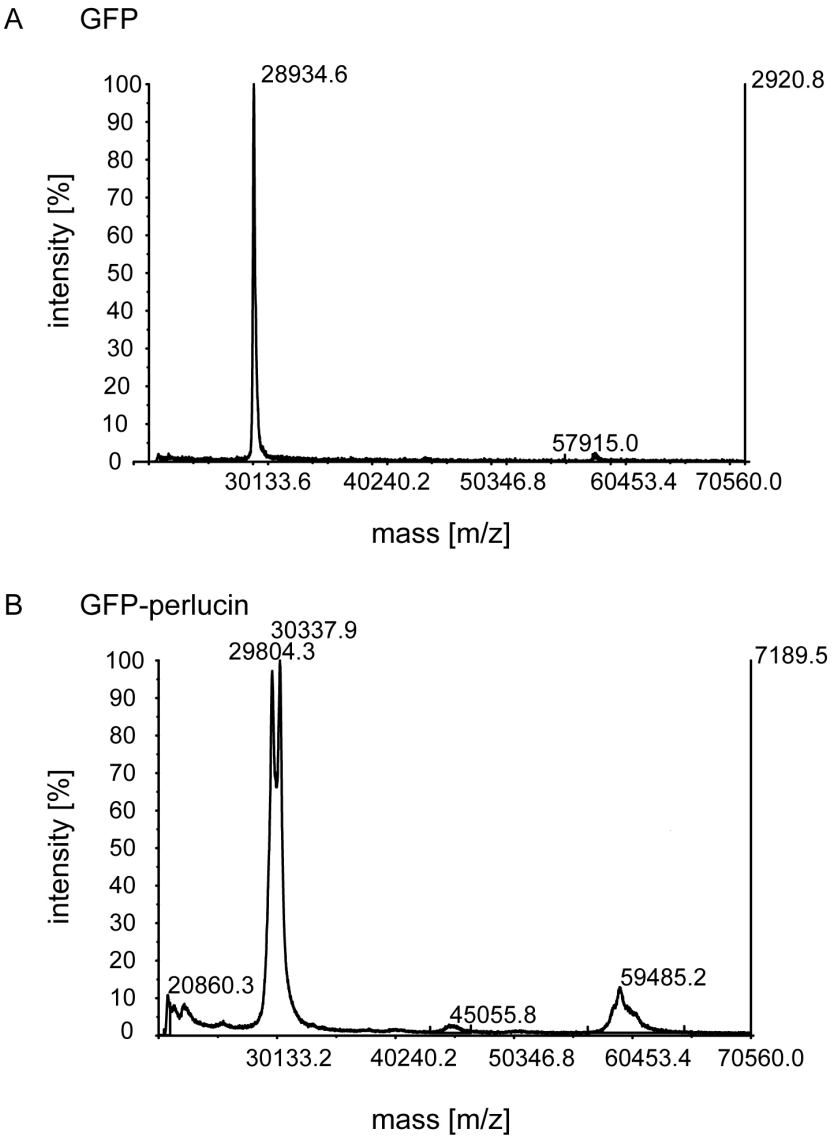
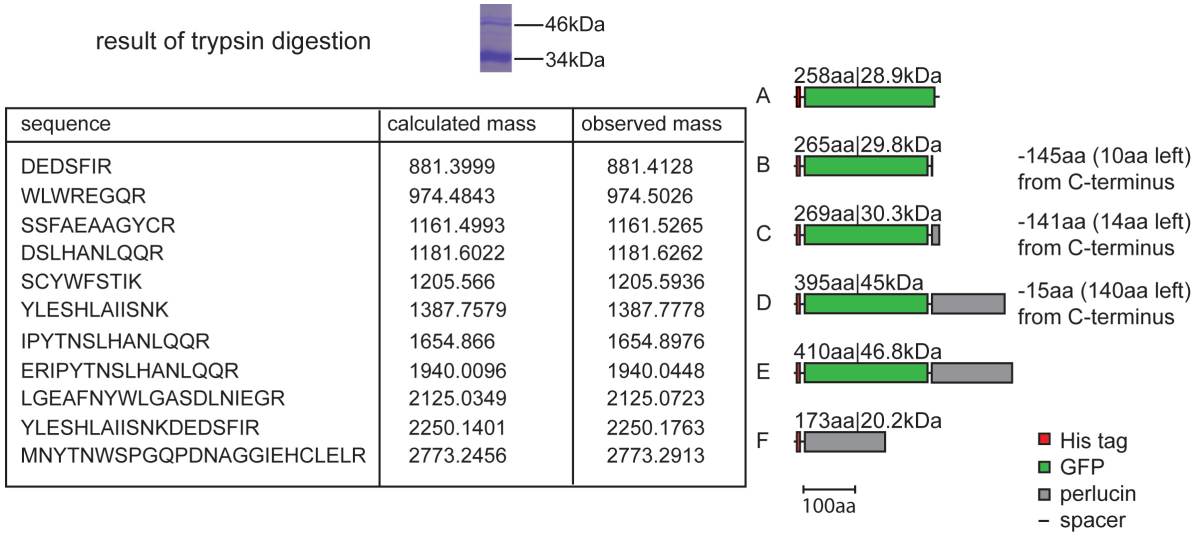


Figure 3.5: Intact mass analysis of Ni-NTA eluates of GFP (A) revealed the expected molecular mass of 28.9 kDa. GFP-perlucin (B) yielded in two major peaks at 29.8 kDa and 30.3 kDa corresponding to truncated protein variants. A minor broad peak at 45 kDa indicated the presence of a 395 amino acid protein variant. Protein dimers indicate additional intact mass peaks in the spectra of GFP (57.9 kDa, A) and GFP-perlucin (59.4 kDa, B).



sequence	calculated mass	observed mass
DEDSFIR	881.3999	881.4128
WLWREGQR	974.4843	974.5026
SSFAEAAGYCR	1161.4993	1161.5265
DSLHANLQQR	1181.6022	1181.6262
SCYWFSTIK	1205.566	1205.5936
YLESHLAIISNK	1387.7579	1387.7778
IPYTNSLHANLQQR	1654.866	1654.8976
ERIPYTNSLHANLQQR	1940.0096	1940.0448
LGEAFNYWLGASDLNIEGR	2125.0349	2125.0723
YLESHLAIISNKDEDSFIR	2250.1401	2250.1763
MNYTNWSPGQPDNAGGIEHCLELR	2773.2456	2773.2913

Figure 3.6: Tryptic digestion of peptides extracted from SDS-PAGE were identified by MALDI-TOF/TOF (left). GFP and truncated GFP-perlucins lacking C-terminal amino acids were identified by either intact mass analysis or after trypsin digestion (right, B-E). Mass fragments of tryptic peptides identified the 46.8 kDa protein as the full length GFP-perlucin including the two C-terminal repeat motifs (left, HANLQQR and right, E).

3.1.3.2 Native-PAGE Analysis of Recombinant Proteins

GFP and GFP-perlucin were separated under non-reducing conditions in the native-PAGE to investigate their agglomeration behaviour. All proteins were visible as fluorescent bands during analysis on the gel. The GFP protein ran in one fluorescent band (Figure 3.7, left A). GFP-perlucin migrated in three different bands in the native-PAGE (Figure 3.7, left B 2-4) and one protein species stucked in the well (Figure 3.7, left B 1). The bands were cut and incubated with equal amounts of 10 mM Tris pH 8.7 for protein extraction and to confirm the presence of the protein variant in a SDS-PAGE with silver staining (Figure 3.7, right A,B). The GFP protein was identified after extraction as indicated in a single band on the SDS-PAGE (Figure 3.7, right A). The analysis of GFP-perlucin bands revealed the co-agglomeration of the 46 kDa fragment with the 29.8-30.3 kDa variant (Figure 3.7, right B 1). The presence of the truncated protein variants in the triple band were verified after protein extraction from the native gel (Figure 3.7, right B 2-4 and Figure 3.6, 29.8-30.3 kDa as confirmed by mass analysis.

The activity of the different protein variants was analyzed in a calcium carbonate precipitation assays as described in result section 3.1.4.4.

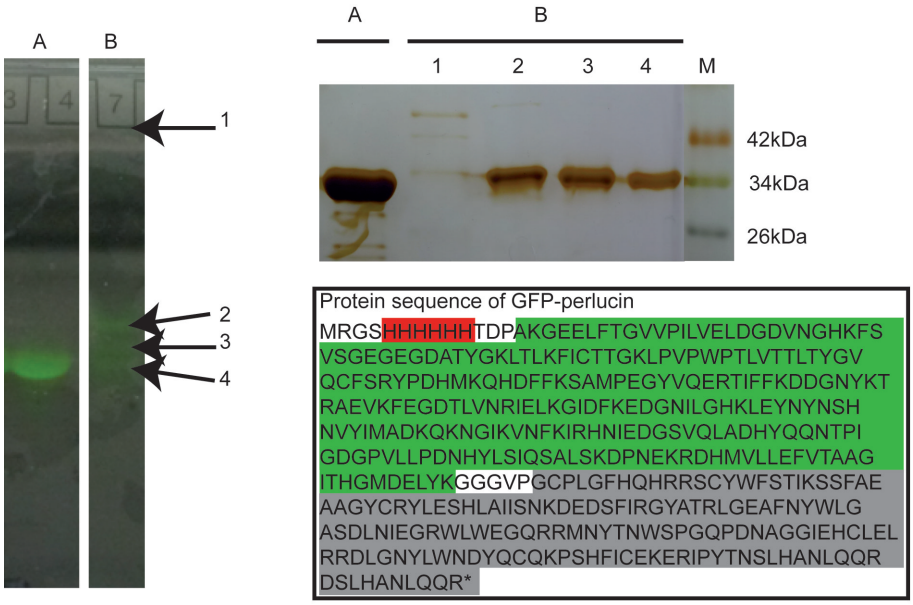


Figure 3.7: Native-PAGE of GFP (28.9 kDa), GFP-Perlucin (46.8 kDa), and SDS-PAGE analysis of extracted GFP-perlucin derivatives. (A) Purified GFP yielded one distinct band. (B) Affinity-purified GFP-perlucin migrated in 3 separate fluorescent bands (2-4) and another fluorescent species remained in the well (1). The latter contained the full-length GFP-perlucin (46.8 kDa) as demonstrated by SDS-PAGE under reducing conditions and silver staining (bottom right, amino acid sequence of perlucin (grey) with affinity tag (red) and GFP (green)). Bands 2-4 were identified as truncated versions of the recombinant protein (34 kDa on the SDS-PAGE contained a 29.8-30.3 kDa fragment as identified by MALDI-TOF/TOF, Figure 3.5). (M) Molecular weight standard, Spectra^{TM} BR.

3.1.3.3 Size Exclusion Chromatography of Recombinant Proteins

For upscaling purposes, size exclusion chromatography was used to separate the GFP-perlucin full length variant from the truncated variants. A representative elution profile was obtained for each protein (GFP and GFP-perlucin) and analyzed by SDS-PAGE as shown for GFP-perlucin in Figure 3.8.

Soluble protein complexes of the full length variant of GFP-perlucin (45-46.8 kDa) and of the truncated variants eluted first (peak A) at a volume comparable to 2000 kDa standard. The truncated variants (B-D, 29.8 kDa and 30.3 kDa) eluted also separately in two peak fractions at elution volumes comparable to the 150 kDa and 29 kDa standards. For control (E), soluble GFP (28.9 kDa) migrated slightly faster on the SDS-PAGE than the truncated variants (B-D) of the GFP-perlucin fusion protein. The result confirmed

the observation from the native-PAGE analysis, where the full length GFP-perlucin variant was not completely separated from the truncated variants (Figure 3.7, right B 1). To analyze the protein activity, two calcium carbonate precipitation assays were performed as described in the next section.

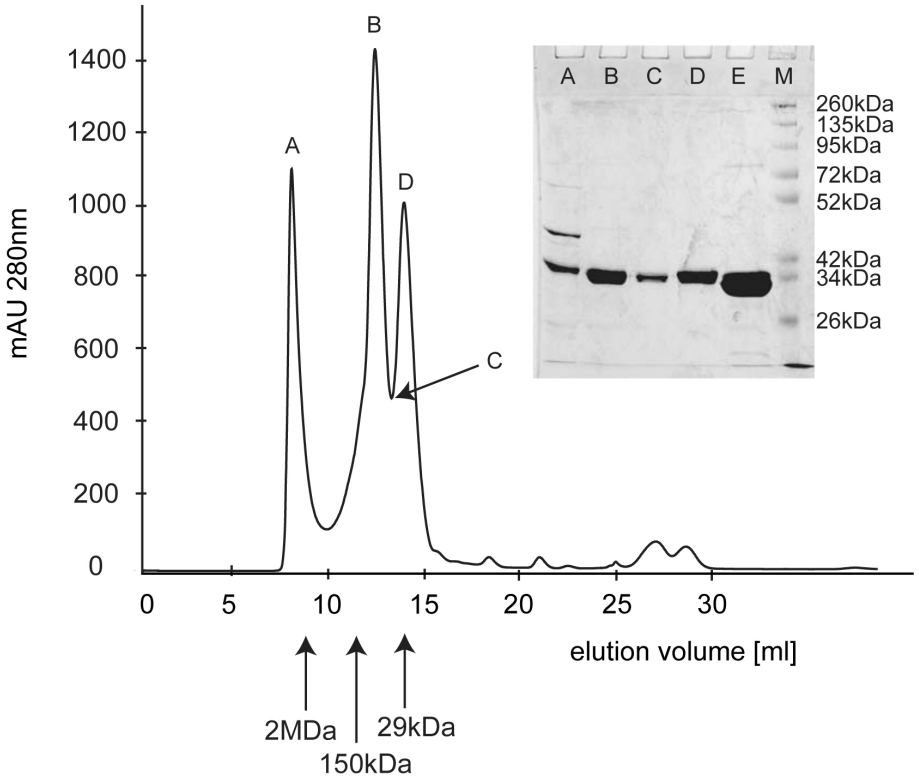


Figure 3.8: An elution profile of a representative GFP-perlucin extract was obtained by size exclusion chromatography and analyzed by SDS-PAGE. Soluble protein complexes of the full length variant of GFP-perlucin (45-46.8 kDa) and of the truncated variants eluted first (peak A) at a elution volume comparable to 2000 kDa standard. (B-D) The truncated variants (29.8 kDa and 30.3 kDa, C) eluted also separately in two separate peak fractions at elution volumes comparable to 150 kDa and 29 kDa standards. (E) Soluble GFP was taken as control, which migrated slightly faster on the SDS-PAGE than the truncated variants (B-D) of the GFP-perlucin fusion protein. (M) Molecular weight standard, Spectra^{TM} BR. Arrows indicate the size standards: Dextran (Ve = 8.9 ml, 2000 kDa); carbonic anhydrase (Ve = 14 ml, 29 kDa), β-amylase (Ve = 11.02 ml, 150 kDa).

3.1.4 Activity Assay of Recombinant Proteins

3.1.4.1 Large Volume Assay and Precipitate Analysis

Analysis of Calcium Carbonate Precipitation Initially, the influence of one-step purified GFP, GFP-perlucin derivatives and a lectin control protein (Concanavalin A) on the precipitation of calcium carbonates was assessed in two different supersaturation assays. In this assay, the rapid pH decrease was recorded over time, after the combination of two precursor solutions ($CaCl_2$ and $NaHCO_3$), whereby the assays (assay I and II) differed in the precursor solution to which the protein eluate was added first (protein-Ca^{2+} and protein-HCO_3^- -precursor). The assay allows to distinguish between additives that acts as inhibitor or promotor of calcium carbonate formation due to a slower or faster precipitation, as compared to the control.

In assay I with protein-Ca^{2+} precursor, one-step purified GFP-perlucin (5 μg, 25 μg, 100 μg), GFP (100 μg), and Concanavalin A (ConA: 25 μg, 100 μg) were diluted in 6 ml $CaCl_2$. When each solution was poured into equal volumes of $NaHCO_3$ solution at pH 8.70 ($\pm$0.11), the pH decreased faster with 5 μg protein (0.4 μg/ml) than with 100 μg (8 μg/ml) to a final pH of 7.50 ($\pm$0.1) (the reaction was completed after 600 sec as observed in all experiments, Figure 3.1, solid line). After 400 sec (Table 3.1), a pH of 7.60 ($\pm$0.2) was achieved. All calcium carbonate suspensions contained the same small-grained, white precipitates (Figure 3.9 vial B).

In the assay II with protein-HCO_3^- precursor, the protein concentration played a major role. At low protein concentrations (5 μg, 0.4 μg/ml) the pH dropped within the first 400 sec, similar to assay I. Surprisingly, at protein concentrations of > 2 μg/ml, pH values remained stable for more than 400 sec (Figure 3.1, dotted line). No visible precipitates were obtained with GFP-perlucin, or GFP (100 μg, 8 μg/ml), once the initial pH shifted from 8.70 ($\pm$0.10) to 8.40 ($\pm$0.07) within 20 sec after combining the two solutions (Table 3.1). Within 2 – 72 h, crystals formed free-floating and strongly attached to the bottom of the glass vial from a clear, transparent solution (Figure 3.9, vial A). In contrast to GFP and GFP-perlucin, the same concentrations of FITC-labelled ConA did not inhibit calcium carbonate precipitation (Table 3.1 for pH values). Controls without protein additive obtained small precipitates as shown for both assays (Figure 3.1, bottom vial C-D).

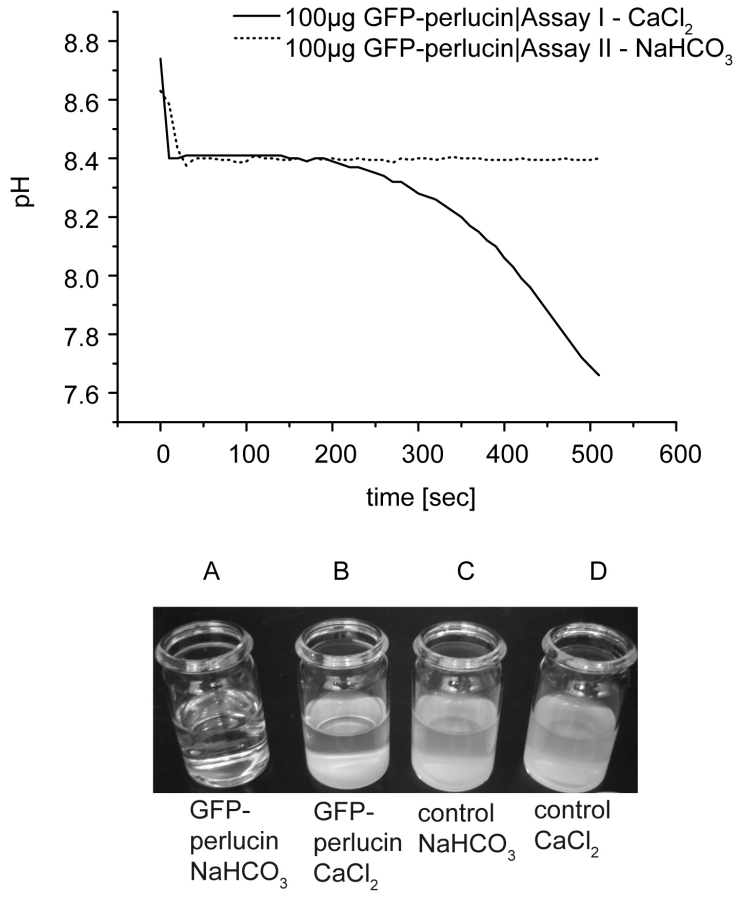


Figure 3.9: $CaCO_3$ precipitation assays I and II. The inhibitory effect of high protein concentrations on the precipitation of calcium carbonate depended on the initial protein precursor solution. The presence of 25 – 100 μg GFP or GFP-perlucin in a $NaHCO_3$ precursor solution delayed the precipitation of calcium carbonate when Ca^{2+} was added to this solution (top, dotted line; bottom, vial A). A protein/$CaCl_2$ precursor solution precipitated within ∼10 min from the addition of HCO_3^- as indicated by the pH curve (top, solid line) and yielded a small grained precipitate (bottom, vial B). While the time-course of the pH drop observed in control experiments without protein additives was similar to the protein/$CaCl_2$ precursor, regardless of which one of the solutions is present in the vial and which one was added, the turbidity of the suspensions (bottom, vials C and D) differed from protein/$CaCl_2$ precursors (bottom, vial B).

Table 3.1: Calcium carbonate precipitation: Effect of protein additives and concentrations. The effect on pH drop was illustrated by selected time points at t= 0 sec, 20 sec and 400 sec. The initial △pH (0-20 sec) was 0.30 ±0.10 for all experiments (n = number or replications). Significant differences became visible only after several minutes. For $NaHCO_3$ as first ionic interaction partner, the second pH shift (△pH 20-400 sec) was lacking after the addition of GFP and GFP-perlucin protein in the amount of 25 μg and 100 μg) (values highlighted in bold). For $CaCl_2$ as the first ionic interaction partner, the second pH shift (△pH 20-400 sec) was reduced compared to experiments with low protein concentration (0.4 μg/ml), or with ConA (2 μg/ml or 8 μg/ml).

Sample	n	pH 0 sec	△pH 0-20 sec	pH 20 sec	△pH 20-400 sec	pH 400 sec
5 μg GFP-perlucin /$CaCl_2$	3	8.84 ± 0.03	-0.3	8.54 ± 0.01	-0.95	7.95 ± 0.03
25 μg GFP-perlucin /$CaCl_2$	5	8.82 ± 0.09	-0.36	8.46 ± 0.05	-0.7	7.76 ± 0.2
100 μg GFP-perlucin /$CaCl_2$	1	8.74	-0.34	8.4	-0.34	8.06
100μg GFP /$CaCl_2$	4	8.52 ± 0.04	-0.26	8.26 ± 0.03	-0.29	7.97 ± 0.22
100 μg ConA /$CaCl_2$	1	8.81	-0.4	8.41	-0.98	7.43
25 μg ConA /$CaCl_2$	3	8.70 ± 0.02	-0.37	8.33 ± 0.01	-0.92	7.41 ± 0.1
5 μg GFP-perlucin /$NaHCO_3$	4	8.79 ± 0.06	-0.3	8.49 ± 0.05	-0.98	7.51 ± 0.05
25 μg GFP-perlucin /$NaHCO_3$	4	8.79 ± 0.11	-0.3	8.49 ± 0.07	**-0.01**	8.48 ± 0.03
100 μg GFP-perlucin /$NaHCO_3$	2	8.63 ± 0.03	-0.2	8.43 ± 0.01	**-0.03**	8.40 ± 0.01
100 μg GFP /$NaHCO_3$	4	8.52 ± 0.0.05	-0.22	8.30 ± 0.03	**-0.01**	8.29 ± 0.02
100μg ConA /$NaHCO_3$	1	8.76	-0.32	8.44	-0.97	7.47
25 μg ConA /$NaHCO_3$	3	8.66 ± 0.02	-0.3	8.36 ± 0.02	-0.93	7.43 ± 0.02

It was demonstrated, that the GFP and GFP-perlucin one step eluates had an inhibitory effect on the formation of calcium carbonate in a concentration of > 2 μg/ml, if supplied in $NaHCO_3$ as first ionic interaction partner. The obtained precipitates were analyzed as described in the next section.

Microscopical Analysis of Calcium Carbonate Precipitates Calcium carbonate precipitates obtained in both assays with one-step purified recombinant GFP-perlucin and GFP, and commercial FITC-labelled ConA were investigated with respect to the most characteristic crystal morphologies and polymorphic types. Minerals obtained from control samples without the addition of proteins are shown in appendix Figure 5.4.

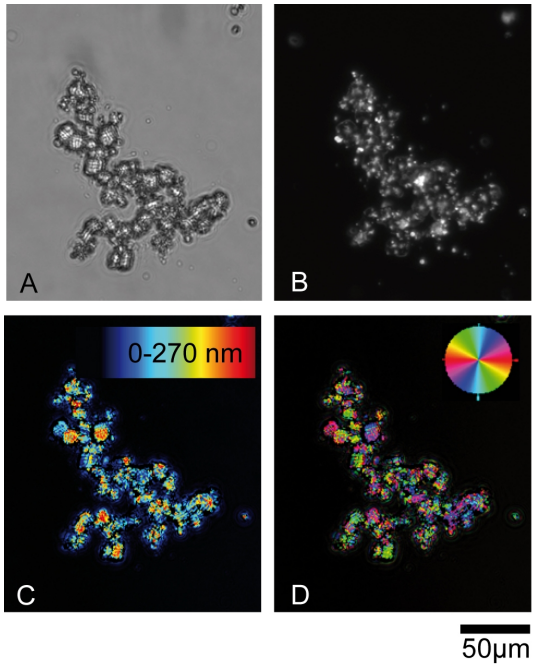


Figure 3.10: Calcium carbonate morphology in assay I (protein/$CaCl_2$ precursor solution). Microscopic analysis of precipitates was performed using bright field (A), fluorescence (B) and LC-PolScope in birefringent retardance (C) and orientation of the slow axis vector (D) imaging modes. The comparative analysis of the micrographs showed that fluorescent GFP-perlucin is closely associated and likely incorporated into spherical calcium carbonate crystals. Fluorescence signals appear inhomogeneous in size and distribution in the centre of mineral particles and slightly weaker on the crystal edges.

Small-grained precipitates were formed within the first 20 sec, if Ca^{2+} was the first ionic interaction partner. These precipitates were either spherical or rhombohedral (Figure

3.10 A, bright field). Fluorescence spots with a broad range of sizes were observed within the majority of crystals, some of them with weakly fluorescent edges (Figure 3.10 B). Most of the precipitates appeared polycrystalline according to LC-PolScope birefringence analysis (Figure 3.10 C,D), which results in a multicoloured pattern in the orientation mode (Figure 3.10 D). Some fluorescent precipitates appeared not or only slightly birefringent. Increased fluorescence was usually associated with spherical structures and disordered agglomerates were favoured (Figure 3.10 B). Non-fluorescent, rhombohedral crystals with rounded edges were obtained in control samples without the addition of proteins (appendix Figure 5.4).

If HCO_3^- was used as first ionic interaction partner, some crystals of several tens of microns in size grown for 3 days from initially inhibited precipitation assays (assay II, 25-100 μg GFP-perlucin (2-8 μg/ml) and 100 μg GFP (8 μg/ml) Figure 3.9, vial A) formed morphologies of a frost pattern type (Figure 3.11 A,B).

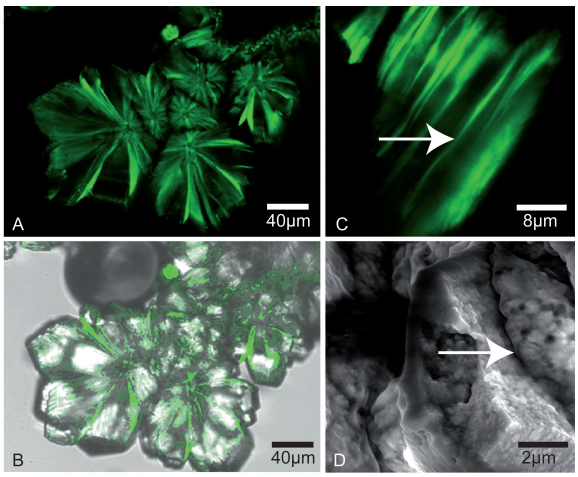


Figure 3.11: Confocal fluorescence laser scanning microscopy images (A-C) demonstrated that GFP-tagged proteins were intimately associated with the mineral phase (B, DIC and fluorescence overlay image) in a layered arrangement of radially expanding (A, B) or lamellar sandwiched (C) aggregates. The mineral platelets were several μm thick and had an internal submicron structure as revealed by scanning electron microscopy (SEM)(D). The platelets were separated from each other by thin, homogeneous layers (arrows) of organic material, which corresponded with the fluorescence signals in (C). SEM image was provided by M. Eder.

Fluorescence and scanning electron microscopic analyses revealed an internal lamellar structure of the composite crystals with a defined arrangement of GFP-proteins preferably at the surface of each lamella. Single lamellae were about 4-5 μm thick and consisted

internally of submicron spherical building blocks (Figure 3.11 D). Fluorescence signals accumulated preferably at the interfaces between the lamellae (Figure 3.11 C). Note that this was the predominant form among variable crystal morphologies obtained in presence of protein. The polymorph analysis with Raman spectroscopy revealed calcite specific spectra for the precipitates in the large volume assay II (not shown).

The structural analysis revealed the interaction of the protein with the mineral phase such that artificial biominerals are produced. At high protein concentration (2-8 μg/ml) of GFP and GFP-perlucin, these biominerals have shown an alternating assembly of mineral-organic layers.

Biochemical Analysis of Calcium Carbonate Precipitates In order to verify the interaction of the mineral and the organic phase, GFP-perlucin was extracted from the biominerals as described in section 2.2.3.11. SDS-PAGE analysis with additional silver staining revealed a clear band on 42 kDa size in the extract sample (Figure 3.12, lane 1), which corresponds to the truncated GFP-perlucin variants. No band was detected in the control lane without protein additive (Figure 3.12, lane 2). The analysis confirmed the direct interaction of GFP-perlucin with the mineral phase as previously observed by fluorescence microscopy analysis.

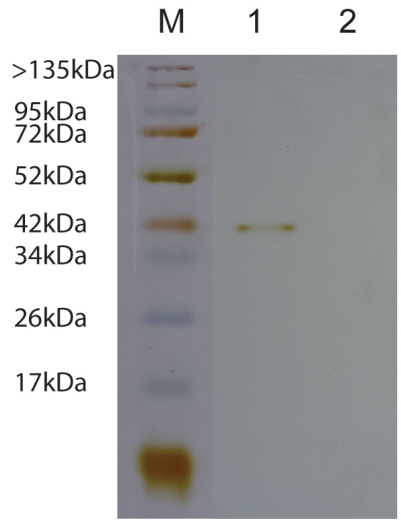


Figure 3.12: SDS-PAGE and silver staining of protein extracts from GFP-perlucin-calcium carbonate composites revealed a distinct band on the size of 42 kDa (2), which migrated above the expected size of 30 kDa and corresponded to the truncated protein variants. No protein band was observed in the control lane without protein (1). (M) Molecular weight marker, Spectra^{TM} BR.

3.1.4.2 Modelling of Calcium-GFP Interaction

To understand the inhibitory influence of high protein concentrations of GFP and GFP-perlucin on the calcium carbonate precipitation in the large volume assay, if HCO_3^- acts as first ionic interaction partner, GFP was analyzed regarding its calcium binding capacity. Potential calcium binding sites were identified with the WebFeature internet tool. The calculation predicted that GFP is able to bind calcium on the top and bottom loop region (loop region I and II) of its β-barrel structure (Figure 3.13, blue dots). A manual analysis of amino acid residues in the protein structure confirmed an accumulation of acidic residues in these two loop regions and revealed two additional binding sites, in the inside of the β-barrel and on its surface. The detailed amino acid analysis can be found in the appendix Figure 5.5.

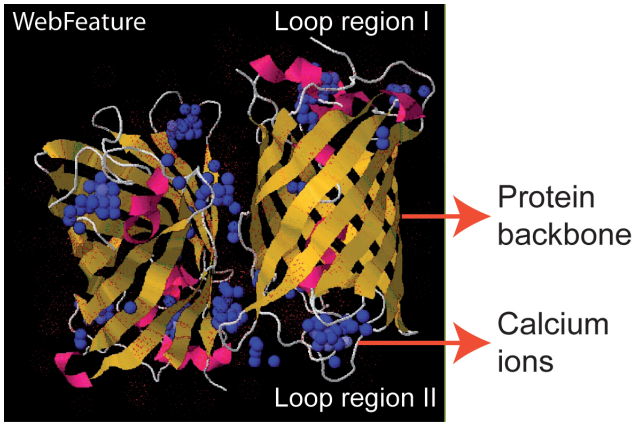


Figure 3.13: Simulation of calcium-binding (blue dots) GFP proteins (wild type GFP model no. 1GFL, [Yang1996]) identified two main binding sites: Loop region I (top) and loop region II (bottom). For a detailed analysis of amino acid composition in the top and bottom regions see appendix Figure 5.5. The prediction was performed with the WebFeature internet tool.

The distances between the acidic amino acid residues (alpha C-atom, calculated with Ballview) were analyzed regarding the calcium ion distances in a typical calcite crystal lattice. The planes of a calcite crystal with a, b = 4.99 Å and c = 10.062 Å would not perfectly fit into the gaps between the acidic amino acids in the loop region I and II. For example, the distance between Asn175 - Gln157 was 5.66 Å and the distance between Gln157 - Asp155 was 5.49 Å (as part of the loop region I, appendix Figure 5.5).

It is noteworthy to mention that the analyzed protein model is based on the wild type GFP (protein model no. 1GFL, [Yang1996]), which differs in 4 amino acid residues from the GFP used in this study [Fuhrmann1999] as identified by sequence alignment

(appendix Figure 5.6). Affected amino acids were Alanine (position 1), Serine (position 2), Phenylalanine (position 64) and Serine (position 65) from wild type GFP, which were exchanged against Proline (position 1), Alanine (position 2), Leucine (position 64) and Threonine (position 65). The exchanged amino acids shared the same physical properties such as hydrophobic residues for the amino acids (positions 1 and 64) and hydrophilic residues (position 65) with the exception of the hydrophilic Serine (position 2), which was changed against a hydrophobic amino acid. Alanine and Serine (positions 1 and 2) were located close to loop region I on the N-terminus of the GFP protein structure and Phenylalanine (position 64) and Serine (position 65) were located in the inside of the barrel structure and were not directly associated to the calcium binding loop region I and II.

3.1.4.3 Field-Flow Fractionation Analysis

For a better understanding of the protein-mineral interaction, a possible ionic strength-dependent agglomeration behavior of one-step protein eluates (GFP and GFP-perlucin) was investigated using field-flow fractionation (FFF). In this work $NaHCO_3$ and $CaCl_2$ were used as precursor solutions to produce $CaCO_3$ crystals as described above.

Since agglomeration of proteins results in an increase of the hydrodynamic radius (Rh), cumulative Rh curves of both protein elutates (GFP and GFP-perlucin) in either $NaHCO_3$ or $CaCl_2$ solution were recorded using the light scattering unit of the FFF device. Addition of protein eluates to $CaCl_2$ solutions led to immediate agglomeration and precipitation of the proteins in the FFF system.

Thus no valid light scattering data of protein eluates in $CaCl_2$ were obtained. However, for $NaHCO_3$ highly reproducible light scattering experiments were performed for both GFP and GFP-perlucin solutions. The Rh of protein eluates in 20 mM $NaHCO_3$ solution was either determined in the presence and absence of Tris-buffer (3 mM). Broader size distributions were calculated for both protein eluates in Tris-buffer free solutions if compared to proteins dissolved in $NaHCO_3$/Tris (Figure 3.14). For GFP-perlucin d50 values of 4 nm ($NaHCO_3$/Tris) and 10 nm ($NaHCO_3$) were calculated. This means that 50 % of the proteins had a Rh of 4 nm ($NaHCO_3$/Tris) and 10 nm ($NaHCO_3$) or smaller. Pure GFP bore d50 values between 2.6 nm ($NaHCO_3$/Tris) and ∼3 nm ($NaHCO_3$) which are smaller than the corresponding GFP-perlucin values.

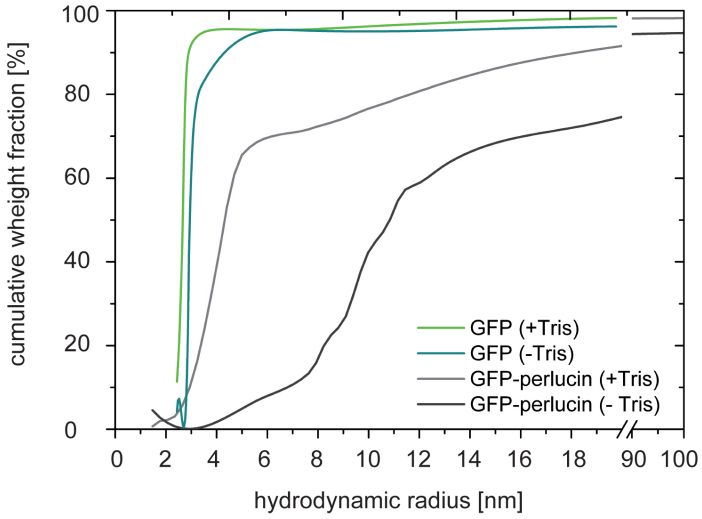


Figure 3.14: Trend lines of cumulative weight fraction analysis of GFP (green lines) and GFP-perlucin (grey lines) one step eluate in 20 mM $NaHCO_3$ pH 8.7 in presence of 3 mM Tris (n = 10, light green ,GFP ;light grey, GFP-perlucin) or without Tris (n = 1, dark green, GFP; dark grey, GFP-perlucin). 50 % of the GFP-perlucin proteins had a Rh of 4 nm ($NaHCO_3$/Tris) and 10 nm ($NaHCO_3$) or smaller. For GFP values between 2.6 nm ($NaHCO_3$/Tris) and ∼3 nm ($NaHCO_3$) were obtained.

3.1.4.4 Small Volume Assay and Precipitate Analysis

Structural Analysis of Precipitates Multiwell plates suitable for microscopy (30 μl) were used in order to assess the role of the reaction volume in GFP and GFP-perlucin induced calcium carbonate precipitations compared to the large volume assay with 12 ml reaction volume. Recombinant GFP and the various GFP-perlucin variants including the 45-46.8 kDa proteins were analyzed individually after extraction from native gels (as described in section 3.1.3.2). Protein concentrations were added to a final concentration of 33 μg/ml, 3.3 μg/ml and 0.33 μg/ml (Figure 3.15) and the remaining precipitates were analyzed microscopically using bright field and fluorescence imaging modes.

The addition of 1 μg (33 μg/ml) of protein yielded in modified crystal shapes in contrast to controls without protein. Recombinant GFP (28.9 kDa) and truncated GFP-perlucin variants (29.8 kDa-30.3 kDa) produced similar or identical fluorescent crystal morphologies (Figure 3.15 A,E,I,M) with tendency to branch (crystals were grown in more than one direction). In both setups fluorescent protein agglomerates were observed in the background.

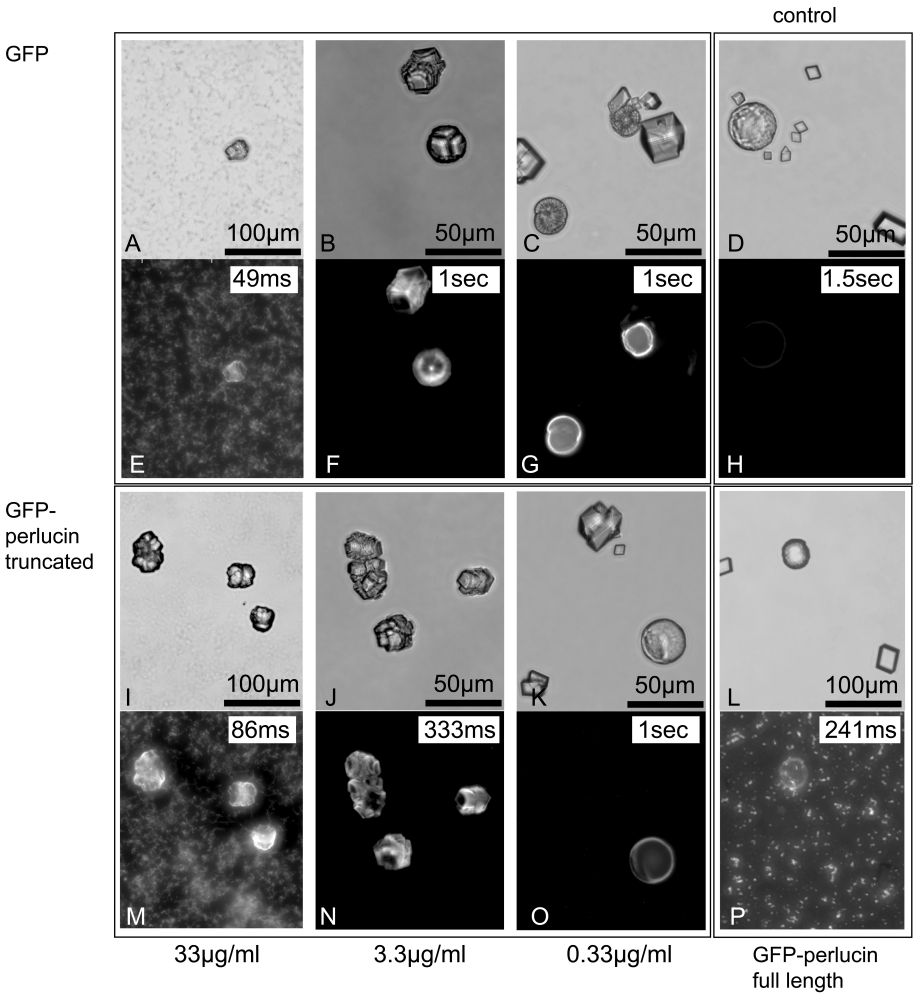


Figure 3.15: Morphologies of precipitates in small volume assays at different protein concentrations. Exclusively branched (crystals prefered to grow in different directions) and elongated fluorescent precipitates were obtained at protein concentrations of 33 μg/μl (A,E,I,M) and 3.3 μg/μl (B,F,J,N). Spherical precipitates obtained with 0.33 μg/μl GFP (C,G) or, with a mixture of all GFP-perlucin variants (K,O) contained fluorescence, in contrast to rhombohedral and octahedral crystals formed under identical conditions in the respective wells. Identical crystal species were obtained without protein (D,H). Precipitates formed in the presence of natively isolated 45-46.8 kDa GFP-perlucin had identical crystal morphologies (L,P) as in the control without proteins (D,H) and multiple fluorescent spots in size ranges of $\sim$1 μm were observed in the background (P).

If 0.1 μg (3.3 μg/ml) was analyzed in the calcium carbonate assay, no fluorescent protein aggregates appeared in the background and the obtained crystal morphologies were similar for GFP and GFP-perlucin. In both samples elongated rhombohedra were the most

prominent crystal shape and the fluorescent signal remained mostly from the elongated crystal sites (Figure 3.15 B,F,J,N).

At very low protein concentrations (0.01 μg; 0.33 μg/ml) and without protein, the most prominent crystal shapes were spheres, rhombohedra and octahedra of individual sizes (Figure 3.15 C,G,K,O for GFP and GFP-perlucin and D,H for control). A difference was observed regarding the fluorescent properties of the spherical crystals, which was strongly fluorescent in the sample with GFP and truncated GFP-perlucin and only slightly fluorescent in the control (Figure 3.15 G,O for GFP and GFP-perlucin and H for control). The residual crystal morphologies with typical calcite geometry had no internal fluorescence properties (Figure 3.15 G,O for GFP and truncated GFP-perlucin and H for control).

When native PAGE extracted GFP-perlucin (45-46.8 kDa) was included in the precipitation assay, typical calcite crystals (non-fluorescent) and slightly fluorescent spheres were observed to co-exist with a strongly fluorescent background of precipitated protein (Figure 3.15 L,P).

In the presence of high concentrations of GFP and truncated GFP-perlucin (1 μg; 33 μg/ml), intensely fluorescent precipitates appeared in the background as described above (Figure 3.15 E,M), reminiscent of the isolated 45-46 kDa GFP-perlucin (Figure 3.15 P). The latter fluorescent background resisted drying the samples with compressed air, whereas the cloudy aggregates of GFP and truncated GFP-perlucin were easily removed and thus were not strongly attached to the microscopy slide.

The small volume assay allowed a fast screening of various protein variants regarding their influence on calcium carbonate precipitation with respect on protein concentration. The properties of the obtained fluorescent mineral spheres were investigated in more detailes as described in the next section.

Raman Analysis of Precipitates Raman spectroscopy was used to identify the polymorph form of the obtained minerals. The spherical particles appeared after 2 hours in controls without protein and in assays with low amounts of protein (<0.01 μg (0.33 μg/ml)) consisted of metastable vaterite with characteristic raman modes at 107 cm^{-1}, 119 cm^{-1}, 270 cm^{-1}, 303 cm^{-1}, 1077 cm^{-1}, 1092 cm^{-1}, slightly shifted 1-3 cm^{-1} to higher wave numbers, in contrast to previous reports [Veinott2009] (Figure 3.16 A). The SEM images bore their rough and layered surface morphology (Figure 3.16 B,C), whereby the residual crystals (rhombohedral, octahedral) were identified as calcite.

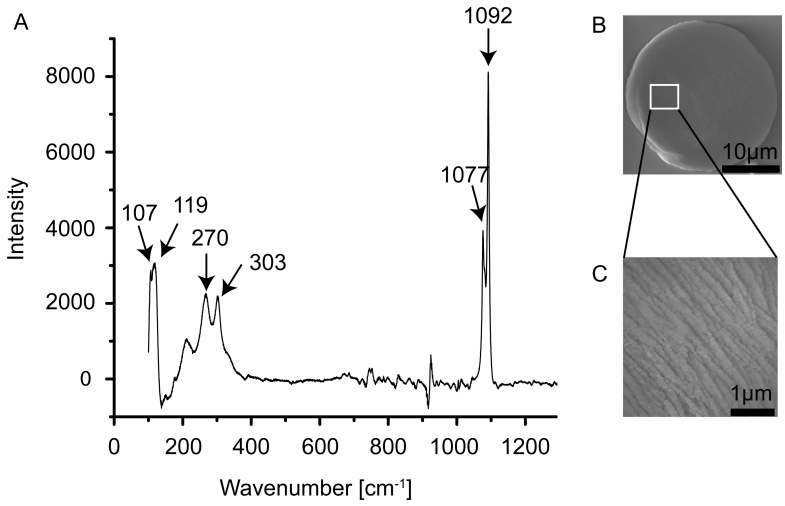


Figure 3.16: (A) Spherical precipitates from small volume assays in the absence or presence ($0.01\mu g/30\mu l$) of GFP derivatives were identified as vaterite with characteristic Raman bands of 107 cm^{-1}, 119 cm^{-1}, 270 cm^{-1}, 303 cm^{-1}, 1077 cm^{-1}, 1092 cm^{-1}. In addition, rhombohedral and octahedral calcite was present. Raman spectra were taken directly in the small volume precipitation chamber, which causes an irregular background. (B,C) SEM images of vaterite bearing their rough and layered surface morphology. Raman spectra were produced by C. Guth and SEM images were provided by M. Eder.

Analysis of Metastable Vaterite Spheres The metastable vaterite particles disappeared after 24-72 hours, with the exact time depending on the particular experimental conditions (Figure 3.17 A-E) in contrast to the calcite mineral phase, which remained stable over time. While vaterite particles were frequently observed in control experiments, fluorescent vaterite particles were observed only when precipitation was performed in the presence of GFP and GFP-perlucin (Figure 3.15 F-J). The metastability of GFP – vaterite composites was followed with time-lapse video microscopy (bright field and fluorescent mode) during the dissolution process. It turned out that after the particles dissolved, an insoluble fluorescent organic skeleton remained stable and attached to the surface from where the mineral originally precipitated (Figure 3.17 J, arrows). However, it was not visible in the transmission light microscopical image (Figure 3.17, E, arrows). Detailed investigation revealed that the dissolution process took place in a characteristic time course (Figure 3.17 A-E). Both, the crystal morphology and the fluorescence intensity distribution changes before the dissolution process was completed. In transmitted light the vaterite particles change their contrast from translucent to dark and again to

translucent during the time lapse experiment.

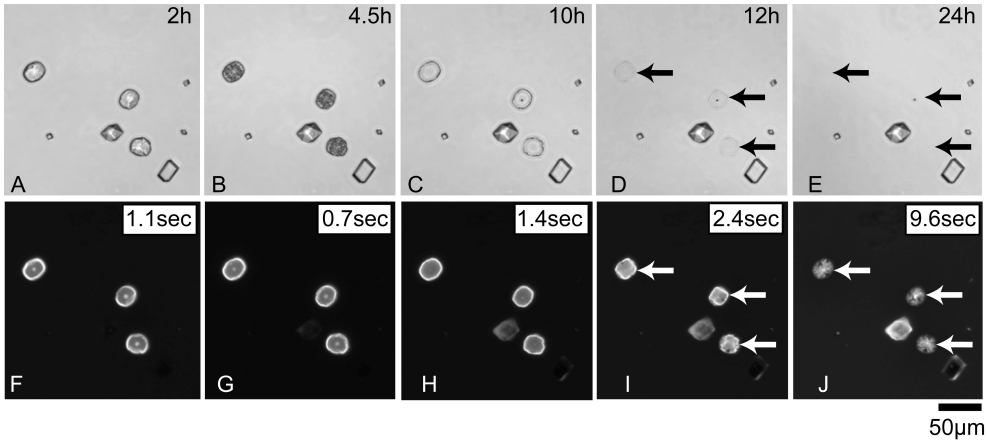


Figure 3.17: (A-E) The spontaneous dissolution of vaterite precipitates was microscopically observed for 12 – 24 h in bright field (A-E) and fluorescence microscopy (F-J) modes. The fluorescent organic matrix associated with the vaterite particles (F-I; arrows) remained partially insoluble after the vaterite disappeared (J, arrows). (A) Both, vaterite spheres and calcite crystals (rhombohedral and octahedral) were visible after 2 hours of precipitation. The vaterite spheres are fluorescent predominantly in their centre and on the surface for up to 4.5 h (F,G). After 12 h, both the morphology and fluorescence intensity of the vaterite particles changed (D,I, arrows). The vaterite particles completely disappeared according to bright field micrographs after 24 h (E), whereas part of the fluorescence signal remained stable (insoluble) at the same position (J, arrows).

A characteristic feature of metastable vaterite spheres was that particles precipitated in the absence of protein had a projected area of 802 μm^2 $\pm$433 μm^2, whereas areas of vaterite coprecipitated with GFP were reduced to 360 μm^2 $\pm$212 μm^2 and with GFP-perlucin to 479 μm^2 $\pm$251 μm^2 (Figure 3.18 A). Note that particle size variations between assays are rather common.

The quantification of the fluorescence intensity (average raw values per pixel of each particle at constant exposure time) showed that vaterite formed in the absence of protein is with 52 $\pm$36 (autofluorescence) significantly lower ($p < 0.001$) as composite particles with GFP fusion proteins (Figure 3.18). GFP – vaterite (152 $\pm$33) and GFP-perlucin – vaterite (119 $\pm$45) have more or less similar protein-mineral ratios, although GFP – vaterite particles appeared slightly more fluorescent (Figure 3.18 B). It was shown, that vaterite spheres are obtained in control experiment and if protein was added in low concentrations. However, the vaterite spheres had gained specific properties due to their

interaction with the organic phase such as shown on the reduced particle area and the remaining insoluble protein patch, which remained after the dissolution of the spheres.

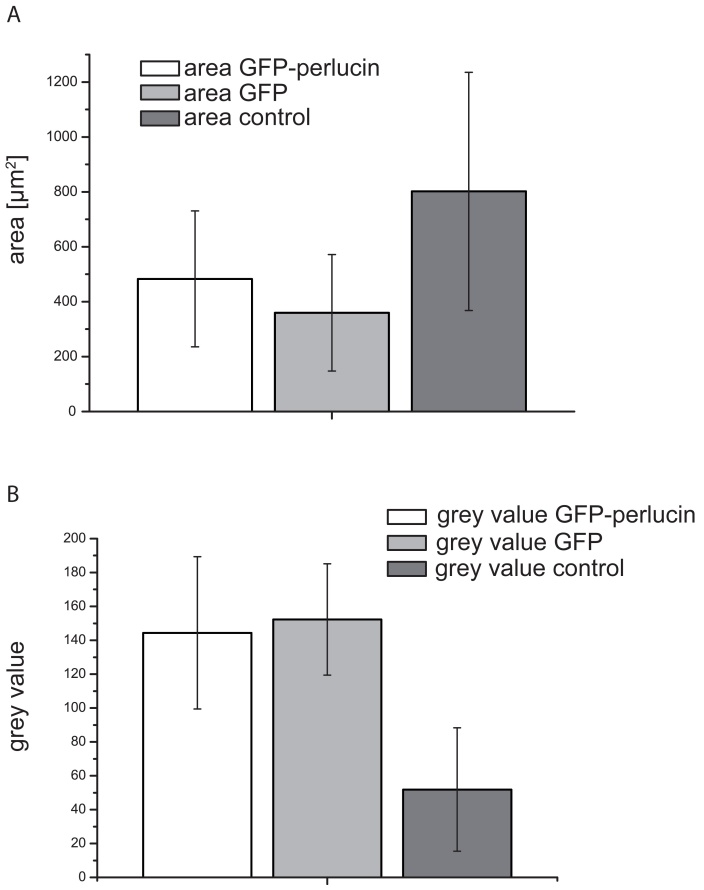


Figure 3.18: (A) Area of the 2-D projection of vaterite particles in bright field microscopy images. (B) Fluorescence of vaterite particle 2-D projections encoded by raw data grey values at constant exposure time of 1000 ms. Only vaterite particles were taken into account. Data shown in (A-C) were calculated with GFP (n = 31), GFP-perlucin (n = 33), control (n = 66). (A), The areas of 2-D projections from GFP and GFP-perlucin vaterite spheres were almost identical (GFP vaterite slightly smaller, $p<0.05$) and with approximately 400 μm^2 (area GFP: 360 μm^2, GFP-perlucin 479 μm^2) about half the size of the protein-free control vaterite (area control: 802 μm^2) and, as such, significantly different ($p<0.001$). (B), Raw fluorescence intensities of vaterite spheres with GFP and GFP-derivatives were approximately 3x higher than the autofluorescence of spherical vaterite formed in the absence of protein ($p<0.001$).

3.1.5 Synthesis of GFP-Silica Nanoparticles

In a cooperation with the Nano Cell Interaction group at the INM (Dr. Annette Kraegeloh), GFP one step eluate was used as basis for the synthesis of GFP-silica nanoparticles (core-GFP, shell-silica) with the aim to produce fluorescent particles for microscopical approaches. The strategy was designed by Dr. Christian Cavelius and the particle synthesis was performed by Sarah Schmidt and Anika Weber. First results demonstrated that the GFP is functionally incoperated into the particle core although leaching effects of the GFP protein had to be overcome in further experiments. A transmission electron microscopical image of the particles with a size of approximately 25 nm is shown in Figure 3.19.

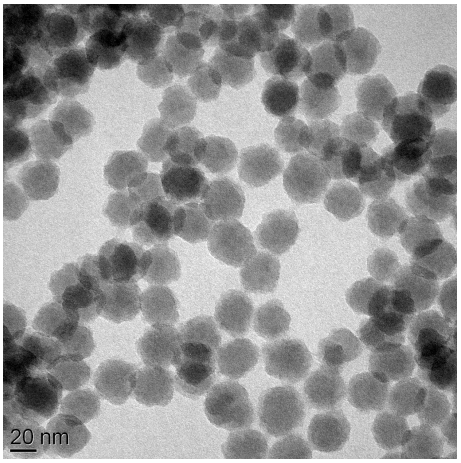


Figure 3.19: Transmission electron microscopy image of GFP-silica particle demonstrates a monomodal distribution with a diameter average of 25 nm. Image was taken by M. Koch.

3.2 Expression of Biomineralization Proteins in Plants

3.2.1 Selection of the Signal Peptides for Cell Wall Delivery

To evaluate the possibilities of the expression of biomineralization proteins in the plant organism, the encoding gene sequences of *perlucin*, *N16N* and *ovocleidin-17* (OC17) were synthesized (Entelechon GmbH) and prepared for the expression in the plant cell wall. The cell wall was used as target compartment, because of its important functions in mechanically supporting the plant cell and mimicking the natural biomineralization matrix in certain aspects.

A fusion of signal peptides deriving from plant cell wall proteins to the N-terminal part of the biomineralization protein were designed with the intention to deliver the biomineralization proteins to the cell wall. Originally, three different signal peptides have been selected to raise the chance that at least one sequence would be functional. The sequences originated from cell wall proteins of *Arabidopsis thaliana*, which are well characterized and associated to three different families of cell wall proteins. Information about the signal peptides and cell wall proteins were obtained from the Arabidopsis tair website (section 2.1.16) and literature.

The most promising candidates were the signal peptides from the cell wall proteins EXT3 (AT1G21310), PRP4 (AT4G38770) and the chitin-hydrolase (AT4G19810). The *EXT3* sequence encodes an extensin, a member of hydroxy proline rich proteins (HPRP) family that plays a major role in cell plate assembly during cell division. The protein is able to form aggregates and can build complex network structures, which provides the mechanical strength to growing cell walls as demonstrated by a lethal knock out mutant of EXT3 [Cannon2008].

The *PRP4* sequence encodes a protein of the proline rich protein family (PRP). This protein family is involved in the first plant defence response, which includes a fast expression and delivery to the infection side of the plant cell wall and subsequent cross-linking due to repetitive elements. The general function of these proteins is discussed to give mechanical strength to primary cell walls. The protein expression of PRP4 in the cell wall was previously confirmed [Fowler1999].

The chitin-hydrolase encodes a chitinase that is a part of glycoside hydrolase family and microarray results suggested that transcript levels rise in response to various stresses such as osmotic stress and pathogen attack. The protein was found to be located in cell wall apoplastic fluid [Boudart2005].

An additional verification of all sequences to encode signal peptides was perfomed with the internet tools iPSORT and SignalP. Please note that iPSORT allows to select plant organisms as host organism for the analysis. A delivery to the chloroplast and mitochondria was excluded for all proteins with the analysis tool iPSORT as shown on the example of the chitin-hydrolase in Figure 3.20. Both analysis revealed the abundance of a signal peptide sequence for all candidate proteins.

iPSORT Prediction

Predicted as: *having a signal peptide*

Sequence (Type: plant)

```
  1 MSSTK LISLI VSITF FLTLQ CSMAQ TVVKA SYWFP ASEFP VTDID SSLFT
 51 HLFCA FADLN SQTNQ VTVSS ANQPK FSTFT QTVQR RNPSV KTLLS IGGGI
101 ADKTA YASMA SNPTS RKSFI DSSIR VARSY GFHGL DLDWE YPSSA TEMTN
151 FGTLL REWRS AVVAE ASSSG KPRLL LAAAV FYSNN YYSVL YPVSA VASSL
201 DWVNL MAYDF YGPGW SRVTG PPAAL FDPSN AGPSG DAGTR SWIQA GLPAK
251 KAVLG FPYYG YAWRL TNANS HSYYA PTTGA AISPD GSIGY GQIRK FIVDN
301 GATTV YNSTV VGDYC YAGTN WIGYD DNQSI VTKVR YAKQR GLLGY FSWHV
351 GADDN SGLSR AASQA WDATT ATTRT IQKV
```

Values used for reasoning

Node	Answer	View	Substring	Value(s)	Plot
1. Signal peptide?	Yes	Average Hydropathy (KYTJ820101)	[6,25]	1.695 (>= 0.9225? Yes)	show
2. Mitochondrial or chloroplast ?	No	Average Negative Charge (FAUJ880112)	[1,30]	0 (< 0.083? Yes)	show
		Indexing: AI1 Pattern: 22121222 (ins/del <= 2)	[1,30]	MSSTKLISLIVSITFFLTLQCSMAQTVVKA 22220212212122222222222222202 221-21222	--
3. Mitochondrial?	No	Average Isoelectric Point (ZIMJ680104)	[1,15]	6.06533 (>= 6.21? No)	show
		Indexing: AI2 Pattern: 100100110 (ins/del <= 3)	[1,15]	-MS-ST-KLISLIVSITF -00-00-10200200200 100100110	--

Figure 3.20: Signal peptide analysis (iPSORT) of the chitin-hydrolase protein sequence (AT4G19810). The analysis identified the N-terminal sequence as signal peptide for the plant secretory pathway and excluded a delivery to the chloroplast or mitochondria.

3.2.2 Cloning Strategy - Change of Resistance Gene

An overview covering the cloning strategy for the expression of biomineralization proteins in the plant cell wall is shown in Figure 3.21. All sequences were delivered in the Gateway entry vector pENTR/D-TOPO, which can be selected in *E. coli* with a kanamycin resistance gene (Km). The pMDC83 vector served as destination vector for the expression in the plant tissue. This vector allows a constitutive expression of the gene of interest and a facilitated protein detection through a fusion to a GFP protein (more vector details are described in the material and method section 2.1.3.2 and in

Figure 2.4). The pMDC83 vector contains a Km resistance gene for selection. To avoid irregularities within the selection procedure, the Km resistance gene was changed in the pENTR/D-TOPO against a gentamycin resistance gene (Gm). The exchange of the resistance gene was repeated for each construct and is described in following part.

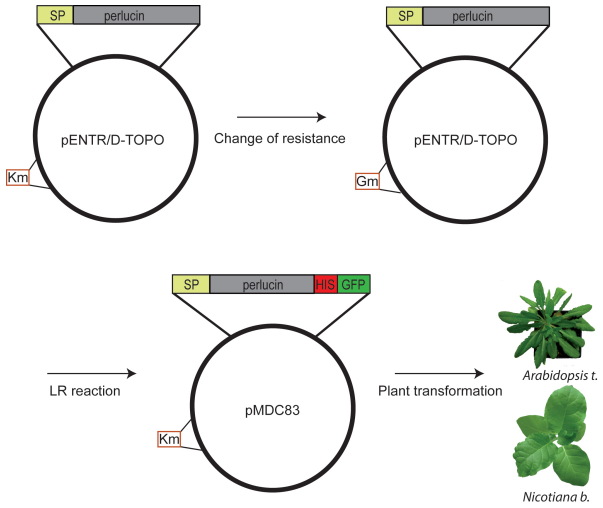


Figure 3.21: Overview about the cloning strategy for the expression of the biomineralization proteins in the plant cell wall on the example of the *perlucin* gene. Initially, the resistance gene in the gateway entry clone pENTR/D-TOPO was changed from kanamycin (Km) to gentamycin (Gm). Otherwise, the selection of the destination vector would either fail or become very difficult, since transformants carrying the entry vector will also survive. In a second step, a LR reaction was performed to transfer the gene of interest into the plant expression vector pMDC83. The pMDC83 plasmid was used for transformation of *Arabidopsis thaliana* and *Nicotiana benthamiana* (tobacco).

The procedure of the resistance gene exchange is described on the example of pENTR/-D-TOPO-SP_{EXT3} -*perlucin*. The pDONR207 vector (contains the Gm resistance gene) and the pENTR/-D-TOPO-SP_{EXT3} -*perlucin* vector were digested with the enzyme *Bsp*HI to excise the resistance genes in both vectors. The restriction digestion reaction mix was loaded on a agarose gel to separate the obtained fragments (Figure 3.22). Fragments of 2199 bp (pENTR/-D-TOPO-SP_{EXT3} -*perlucin*-vector) and 927 bp (Km resistance gene) were obtained (Figure 3.22 A) and the 2199 bp fragment was isolated from the agarose gel. The undigested vector (3126 bp) migrated in one prominent band on a size of 2000 bp due to its circular structure. The digestion of the pDONR207 vector resulted in two fragments with the size of 3543 bp and 2042 bp (Figure 3.22, C). The 2042 bp fragment contained the Gm resistance gene. This fragment was isolated

from the agarose gel and ligated with the 2199 bp fragment from the former digestion. The ligation reaction was transformed into *E. coli* cells and plated on LB agar plates containing Gm. Growing colonies were selected for further analysis.

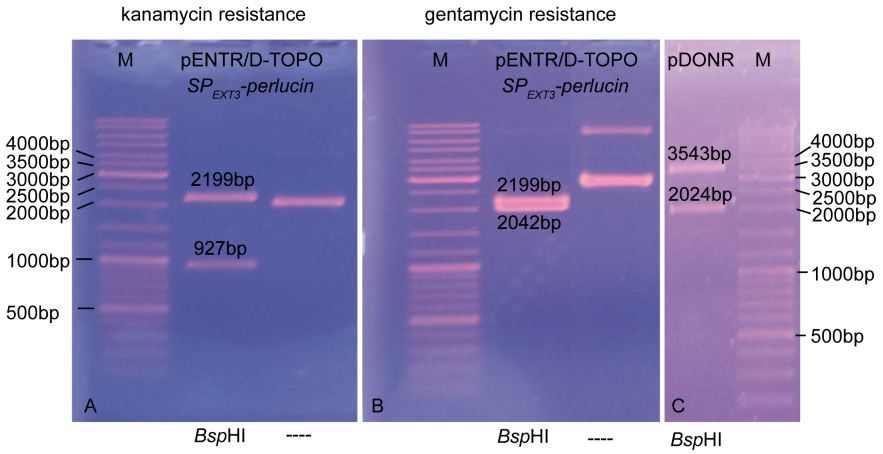


Figure 3.22: Restriction digestion of pENTR/-D-TOPO-SP_{EXT3}-*perlucin* and pDONR207 with *Bsp*HI separated on a agarose gel. Digestion before the exchange of resistance gene revealed two fragments of 2199 bp (pENTR/-D-TOPO-SP_{EXT3}-*perlucin*, vector) and 927 bp (Km resistance gene) (A). After the exchange, the obtained fragment sizes were 2199 bp (vector) and 2042 bp (Gm resistance gene, B lane 2). The Gm resistance gene originated from the pDONR207, shown after *Bsp*HI digestion with fragments of 3543 bp (vector) and 2042 bp (Gm resistance gene), (C). (M) Molecular weight standard, GeneRuler^{TM}.

The presence of the Gm resistance gene was analyzed by PCR reaction using the gentamycin resistance primers. The PCR reaction was positive for 7 clones. The plasmids were isolated from two clones and digested similar as described above. The gel image revealed that the change of resistance was successful. There was one band of 2199 bp (pENTR/-D-TOPO-SP_{EXT3} -*perlucin*-vector backbone) and one of 2042 bp fragment (Gm resistance gene), as expected. The undigested vector with a size of 4241 bp ran in two distinct bands due to its variable secondary structures (Figure 3.22, B).

The LR reaction was performed to transfer the gene of interest into the pMDC83 plant expression vector (plasmid map, Figure 2.1.3.2). The reaction was transformed into *E. coli* cells and positive clones were verified by PCR reaction using gene specific primers. Subsequently, plasmids were isolated from positive selected clones and the inserted gene was sequenced before plant transformation.

3.2.3 Expression of Perlucin-GFP in Arabidopsis thaliana

To investigate the possibility of the expression of biomineralization proteins in plants including the selection of transgenic lines, the plant model organism *Arabidopsis thaliana* was used to express perlucin (called perlucin-GFP). Perlucin was chosen as first candidate for the reason to be the most promising protein for promotion of calcium carbonate formation *in vivo* and *in vitro* (described in the introduction). Moreover, it has lectin properties, which would match closely to bind sugar residues in the plant cell wall and define its position. For possible microscopic *in vivo* detection, all proteins were fused C-terminally to a GFP (part of the expression cassette in the pMDC83). The *perlucin-GPF* gene was transformed into *Arabidopsis thaliana*, which where then selected and analyzed. The transformation procedure is described in material and methods section 2.2.4.7. The two signal peptide sequences *EXT3* and *PRP4* were selected for *perlucin-GFP* transformation of *Arabidopsis thaliana* after their integration into the pMDC83 vector.

3.2.3.1 Hygromycin Selection of T0 Plants

Seeds from transformed plants (t0) were harvested and grown on MS agar containing hygromycin for the selection of positively transformed seedlings as shown in Figure 3.23.

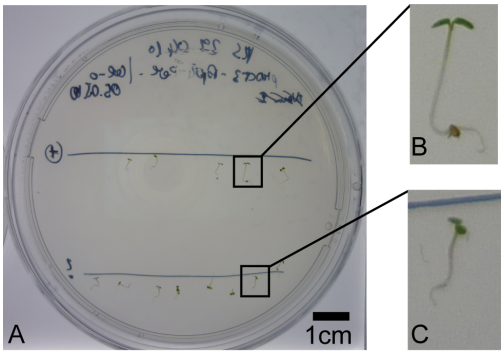


Figure 3.23: Hygromycin was used as selection marker to distinguish *perlucin-GFP* transformed *Arabidopsis thaliana* plants from untransformed plants. The positive plants are elongated (A, top row and B) compared to untransformed plants (A, bottom row and C).

The elongated plants were selected and transferred to the soil in order to allow them growing to the adult stage and building seeds. Each selected plant was genotyped on DNA level to confirm the presence of the *perlucin-GFP* gene as described in the next

section. Subsequently, the selected plant lines were enumerated with the initials of the experimentator (EW).

3.2.3.2 Analysis of the T1 Plants

Genotyping of T1 Plants To confirm the insertion of biomineralization gene sequences into the plant genome, plants were genotyped (Table 3.2). Only plants, which revealed 2-3 positive signals in the PCR reaction with the different primer combinations (Figure 2.12) were used for further analysis (Table 3.2, bold).

Table 3.2: Summarized genotyping results from *perlucin-GFP* transformed *Arabidopsis thaliana* (t1) plant material. Different primers were used to select the positively transformed plants (Figure 2.12, (x) represent a band in the PCR, (-) no band in the PCR). Selected plant lines used for further analysis are highlighted in bold.

Line	pMDC83-*SP-perlucin*	2x35s	Perlucin	SP-perlucin	SP-GFP	GFP
EW5	SP_{EXT3}	x	-	-	x	-
EW6	SP_{EXT3}	x	-	-	-	x
EW7	SP_{EXT3}	x	x	-	-	x
EW16	SP_{EXT3}	-	x	x	x	x
EW17	SP_{EXT3}	x	x	-	-	x
EW8	SP_{PRP4}	x	-	-	-	x
EW9	SP_{PRP4}	x	-	-	-	x
EW10	SP_{PRP4}	x	-	x	-	x
EW11	SP_{PRP4}	x	-	-	-	x
EW12	SP_{PRP4}	x	-	x	-	-
EW13	SP_{PRP4}	x	-	x	-	-
EW14	SP_{PRP4}	x	-	x	-	-
EW15	SP_{PRP4}	-	-	x	-	x
EW18	SP_{PRP4}	x	-	-	-	-
EW19	SP_{PRP4}	x	-	-	-	x
EW20	SP_{PRP4}	x	x	-	-	x
EW21	SP_{PRP4}	x	-	x	-	x
EW22	SP_{PRP4}	x	-	x	-	x
EW24	SP_{PRP4}	x	-	-	-	x
EW25	SP_{PRP4}	-	-	x	-	-
EW26	SP_{PRP4}	x	-	-	-	x

Due to the observation of unspecific bands obtained with GFP and 2xCaMV35s specific primers, SP-perlucin primers were used for further genotyping analysis (combination (3) in Figure 2.12). Four different lines were selected for each signal peptide. The lines EW5, EW7, EW16, EW17 contained the *EXT3* signal peptide gene sequence and the lines EW10, EW20, EW21, EW22 contained the *PRP4* signal peptide sequence. The lines were grown to the t3 generation (material and method section 2.2.4.9) and the t2 and t3 generations were analyzed. To ensure the correctness of the insert, PCR bands were excised from the agarose gel and the DNA was extracted for sequencing, which confirmed the presence of the *perlucin* gene.

3.2.3.3 Analysis of the t2 Generation

Germination Rate and Root Growth Assay of T2 Plants T2 plants were grown on agar plates for 12 days and the germination rate was determined. Whereas the wild type (wt) germinated to 88 %, the EW7 line had a significantly minor germination rate of 32 %. In addition, the germination rate of EW20 and EW22 was reduced to 74 % and 68 %. The lines EW5, EW10, EW16, EW17 and EW21 had a germination rate of about 81-92 % (appendix Table 5.4). Regarding the sensitivity of the plant root system due to a modified gene transcription, the mean root length of selected perlucin-GFP plants was analyzed (Figure 3.24).

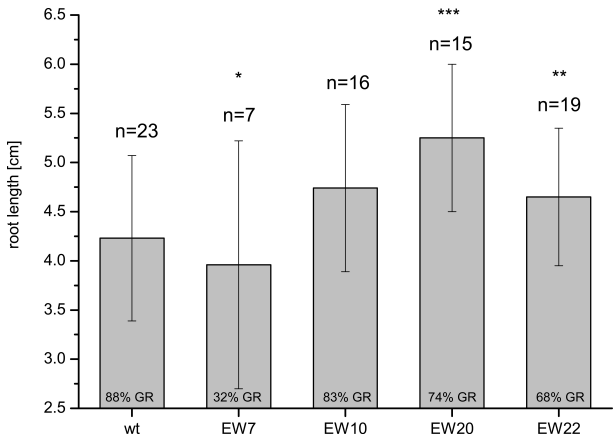


Figure 3.24: Analysis of mean root length of wild type (wt), EW7, EW10, EW20 and EW22 line. EW20 ($p<0.001$ (***)) and EW22 ($p<0.01$ (**)) plants had a significant increased mean root length with 5.25 cm and 4.65 cm, respectively, compared to wt plants with 4.23 cm. EW7 ($p<0.05$ (*)) and EW22 had a less significant difference. Error bars indicate the standard deviation, (n) number of analyzed plants, (% GR) percent of germination rate.

The obtained mean root lengths were: EW7 (3.96 ±1.26 cm), EW10 (4.74 ±0.85 cm), EW20 (5.25 ±0.75 cm), EW22 (4.65 ±0.70 cm) and wt (4.23 ±0.84 cm). The EW10, EW20, EW22 and wt plants had a mean root length in the range of 4.25 to 5.25 cm. The EW20 line had the highest mean root length of 5.25 ±0.75 cm and a deviation of 1 cm compared to wt plants. The shortest mean root length was in the EW7 line with 3.96 ±1.26 cm. Statistical analyses revealed that the wt root length is slightly higher compared to the EW7 ($p<0.05$ (*)), EW20 ($p<0.001$ (***)) and EW22 ($p<0.01$ (**)) lines. Statistical investigations showed no significant difference between the wt and the EW10 line. Images of the plants on agar plates are shown in appendix Figure 5.10. The results showed that the EW7 line had a minor germination rate and a slightly shortened mean root length, whereby the EW20 and EW22 plants had a reduced germination rate and an increased mean root length.

Structural Analysis of the T1-T2 Plant Material Seeds were analyzed with scanning electron microscopy imaging (SEM) to investigate their morphology due to the reduced germination rate observed for EW7, EW20 and EW22 line. Examples of wt and EW7 seeds are presented in SEM images (Figure 3.25). Please note that the seed envelope originates form the t1 plants and that only the embryo has a t2 genotype.

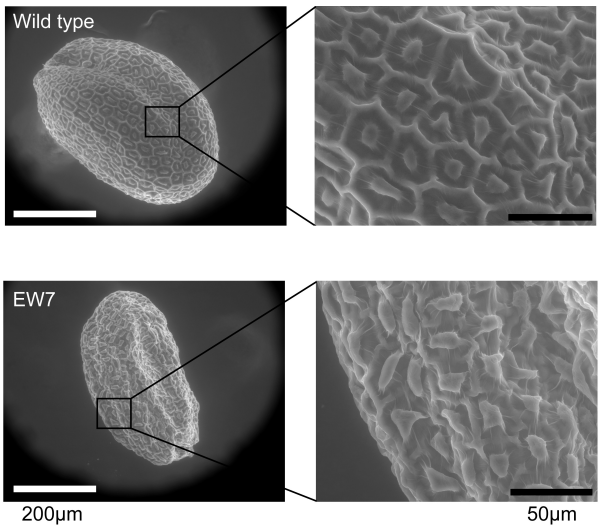


Figure 3.25: SEM images of wt seeds and seeds containing the *perlucin-GFP* gene sequence. The wt seed (top row) had a well defined honeycomb structure compared to the EW7 seed (bottom row). Images were kindly provided by M. Eder.

A deformation of the seed surface was found for seeds of EW7, EW20, EW21 and

EW22 lines compared to wt seeds. The analyzed wt seeds were round and honeycomb structured with an elevation in the middle of each honeycomb. For comparison, the seeds of the EW7 line appeared more sensitive to drying and had no clear honeycomb structure around each elevation. The EW20, EW21, EW22 seeds had only slightly reduced honeycomb structures on their surface compared to wt seeds.

Expression Analysis of T2 Plants at the RNA Level In order to correlate the phenotypical observation with cellular activities, the gene expression level of *perlucin-GFP* lines was analyzed for a pool of t2 generation plants with quantitative real time PCR (qPCR). All constructs contained a 2xCaMV35s promotor for a constitutive gene expression and a high expression level would be expected. Due to the limited amount of plants, the analysis was performed only for selected lines (EW7, EW10, EW20, EW21 and wt).

First, RNA was extracted from a pool of *perlucin-GFP* transformed t2 plants. A representative agarose gel with RNA is shown in the appendix Figure 5.9. Subsequently, the RNA extracts were treated with DNAse before cDNA synthesis. Quantitative PCR analysis was performed after cDNA synthesis. The success of the cDNA synthesis was investigated with control primers for the housekeeping gene *EF1Balpha2*. The PCR reaction was positive for the control reaction, indicating an intact cDNA. An absolute quantification method was used to determine the amount of copies with a standard curve. The standards for the *perlucin-GFP* gene were prepared according to the material and method section and the optimal primer concentrations were determined in a primer matrix reaction before starting experiments with plant samples. The efficiency of the reaction was 1.06 (1 as optimal value), indicating a good primer binding capacity. The standard curve involving 6 points and a linear regression curve (with a coefficient of determination of $R^2 = 0.996$) is shown in Figure 3.26, top. The *perlucin-GFP* cDNA samples and the controls achieved a signal in the range of the lowest standard (Std 10^2 = 100 DNA copies). It would be expected that a gene expressed with the 2xCaMV35s promotor gives a signal between standards containing 10^6 - 10^5 DNA copies. The melting curve analysis revealed one single fragment at 81 °C in all samples that means there was no by-product produced in the PCR reaction (Figure 3.26, bottom).

In contrast to the expected result, no expression of *perlucin-GFP* was detectable in the analyzed plant lines with qPCR. The reaction was repeated 4 times with identical results. The usage of different polymerases had no influence on the reaction.

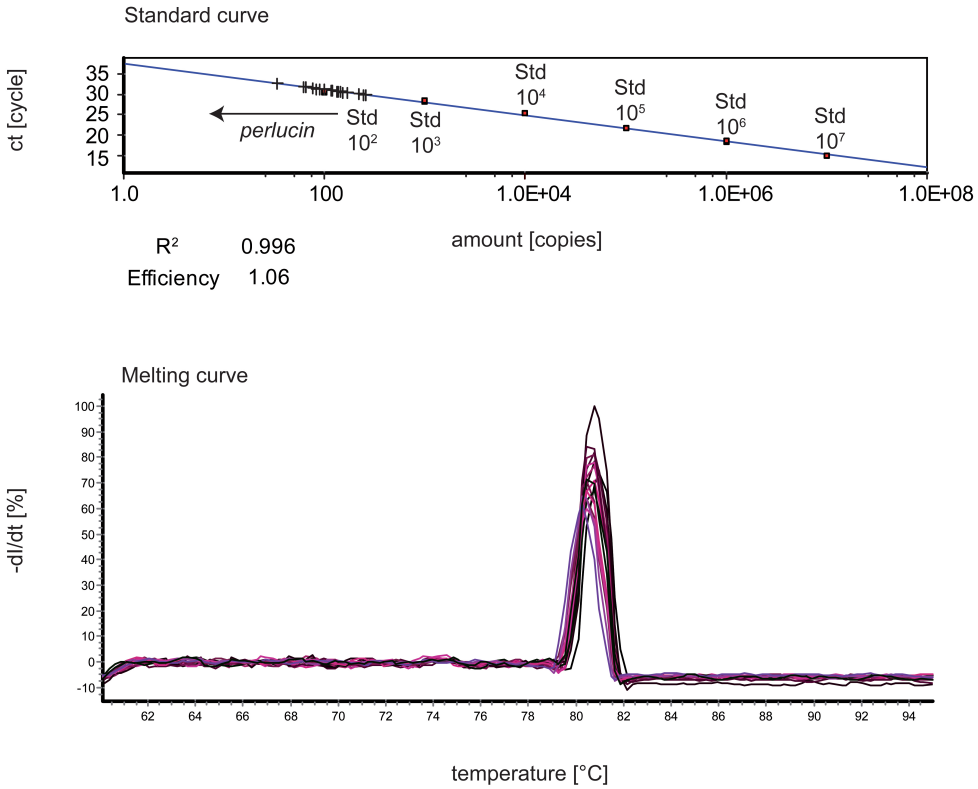


Figure 3.26: Result of qPCR analysis of *perlucin-GFP* transformed *Arabidopsis thaliana* plants. Standard curve involving 6 points and a linear regression curve (R^2 = 0.996, top). The *perlucin-GFP* cDNA samples and the controls achieved a signal in the range of the lowest standard (Std 10^2), indicating that the expression is probably under the detection limit. The melting curve revealed a single fragment at 81 °C in all samples (bottom).

Western Blot Analysis of the T2 Plant Material Despite the result that no transcript of *perlucin-GFP* was observed, the occurence of heterologously expressed proteins was tested by SDS-PAGE and Western blot analyses. Proteins were extracted from *Arabidopsis* leaves and separated by SDS-PAGE. Western blot analyses with an anti-GFP antibody revealed no detectable signal for the perlucin-GFP protein (data not shown). However, the plants were selected further by PCR based genotyping and the t3 generation was analyzed as described in the next section.

3.2.3.4 Analysis of the T3 Plants

Genotyping and Germination Rate of the T3 Perlucin-GFP Plants The t3 generation was genotyped to search for the homozygous *perlucin-GFP* plants from each line.

The progeny will be positive for the *perlucin-GFP* gene, if the mother plant was homozygous (compare also material and method section 2.2.4.9). Therefore, 8-10 independent progeny plants from one mother plant were analyzed. Subsequently, 50 seeds of each single progeny plant were grown on agar plates to determine the germination rate and 10 plants (out of the 50) were used for genotyping. The result of the t3 genotyping is listed in Table 3.3. No positive plants were detected for the EW10 and EW20 lines. For the residual lines a minor number of positive plants were selected (<5 to 16 %). The germination rate varied among each line as shown for EW22 in between 5-100 %, Table 3.3. For example, a detailed genotyping result of EW7 is shown in the appendix Table 5.5. There was no correlation between germination rate and positive result for the inserted *perlucin-GFP* gen. Positive plants were obtained from plants with 100 % germination rate as well as from plants with only 40 % of germination rate as shown for the EW7 line in appendix Figure 5.5 and 5.11.

Table 3.3: Gentotyping result and germination rate of the t3 generation of *perlucin-GFP* transformed *Arabidopsis thaliana* plants. Low numbers of positive plants were identified (<5-16 %).

Line	Number of plants	Positives	Positives	Germination rate
			[%]	[%]
EW5	80	13	16	19-100
EW7	80	3	<5	40-100
EW10	80	0	<5	57-100
EW16	80	3	<5	65-100
EW17	80	1	<5	86-100
EW20	80	0	<5	17-100
EW21	80	9	11	14-100
EW22	80	3	<5	5-100

Phenotype Analysis of T3 Plants The positively selected plants from the t3 genotyping were analysed regarding their phenotype. Figure 3.27 A showed a typical phenotype of a EW7 plant compared to wt. Furthermore, stereo microscopical analysis revealed that the siliques resemble wt in size and form (Figure 3.27 B), but contained empty positions without seeds (Figure 3.27 B, arrows). The rosette leaves revealed wt shape (Figure 3.27 C) and the flowers had a correct structure with 4 petals (white), 6 stamina

and one ovary (Figure 3.27 D).

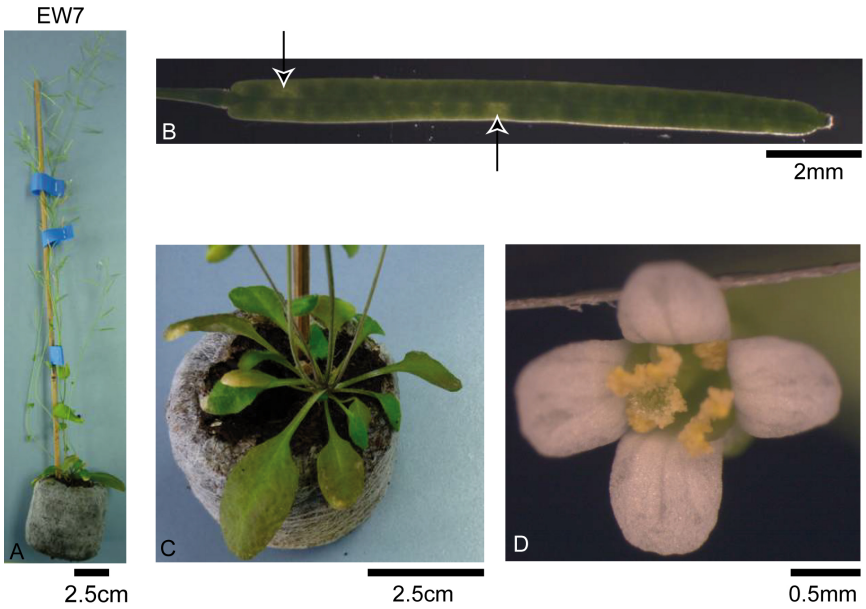


Figure 3.27: EW7 plant revealed wt phenotype after the growth on soil (A-D). Siliques had wt length and form (B) with free positions in the siliques (B, arrows). Rosette leaves (C) and flowers (D) showed the same morphology as wt plants with 4 petals (white), 6 stamina (yellow) and one ovary.

Due to the reduced germination rate, the siliques were analyzed regarding the empty positions without seed (Figure 3.27 B, arrows). The relative number of empty seed positions was determined by dividing the number of empty positions through the number of total positions (Figure 3.28). Absolute values of silique length, number of seed positions and empty positions used for the calculation are shown in appendix Figure 5.12.

There is a tendency to a reduced seed formation in the lines EW16 ($p<0.05$ (*)), EW17, EW21 ($p<0.05$ (*)) and EW22 compared to wt (Figure 3.28). EW16 and EW21 showed a mean reduction of 25-30 %. Please note that siliques were found with normal as well as with reduced number of seeds in all lines. No differences were observed for the EW5 and EW7 lines although the EW7 lines bore the strongest reduction in the germination rate before.

Furthermore, the length of the siliques was measured. There were no differences between the perlucin-GFP and the wt plants regarding silique length, which have been in the range of 8 to 14 cm as shown in appendix Figure 5.12.

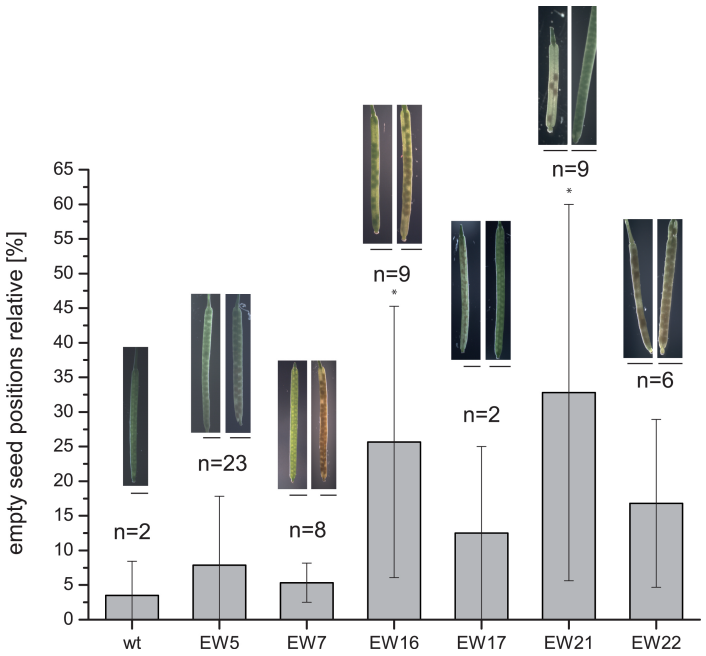


Figure 3.28: Relative number of empty seed positions in the siliques. Significant reduction in number of seeds was observed for the EW16 ($p<0.05$ (*)) and EW21 ($p<0.05$ (*)) lines with 25-30 % compared to wt plants. No significant change was observed for EW5, EW7, EW17 and EW22 line. Error bars indicate the standard deviation, (n) number of analyzed siliques. Bar = 2 mm

3.2.4 Summary: Expression of Perlucin-GFP in Arabidopsis thaliana

T2 and t3 plants of EW5, EW7, EW10, EW16, EW17, EW20, EW21 and EW22 were investigated regarding phenotype, gene- and protein expression of perlucin-GFP (Table 3.4). In summary, the transformation procedure was successful as shown by the positive genotyping results, whereby neither RNA (Figure 3.26) nor protein was detected in the selected t2 lines. Seed morphological changes were observed in t2 siliques of the lines EW7, EW16, EW21, EW22 as shown by SEM images (Figure 3.25, EW7 for example) whereas the residual phenotype of t3 plants had wt appearance (Figure 3.27). Furthermore, a reduction in germination rate was obtained for EW7, EW20 and EW21 plants (t2 generation, appendix Table 5.4 and t3 generation Table 3.3). There was no correlation between reduction in germination rate and the number of seeds in the siliques with exception of the EW21 line. Futhermore, PCR results revealed a reduction of positive plants in the t3 generation compared to the t2 plants as illustrated in detail in Tables 3.3 for the t3 generation.

Table 3.4: Summarized analysis of *perlucin-GFP* transformed *Arabidopsis thaliana* plants compared to wild type plants (wt). (+) Positive, (-) negative, (↓) reduction, (↑) increase, (*) $p<0.05$, (**) $p<0.01$, (***) $p<0.001$, (n.a.) not analyzed because of a limited number of seeds.

Line	PCR result t2	Root length t2	RNA t2	Germination rate t2	Seed morphology t2	PCR result t3	Siliques t3
EW5	+	n.a.	-	wt	n.a.	+	wt
EW7	+	↓ (*)	-	↓	+	+	wt
EW10	+	n.a.	-	wt	n.a.	-	-
EW16	+	n.a.	-	wt	n.a.	+	↓ (*)
EW17	+	n.a.	-	wt	n.a.	-	-
EW20	+	↑ (***)	-	↓	+	-	-
EW21	+	n.a.	-	↓	+	+	↓ (*)
EW22	+	↑ (**)	-	wt	+	+	wt

3.2.5 Expression of Biomineralization Proteins in Nicotiana benthamiana

Based of the findings from the analyses of *perlucin-GFP* transformed *Arabidopsis thaliana* plants, where neither RNA nor protein was detected so far, N16N (introduction 1.1.4.2) and ovocleidin-17 (introduction 1.1.4.3) have been targeted for the expression in the plant organism. Identical cloning strategy was used as described for perlucin (result section 3.2.2). Because *Arabidopsis thaliana* is very time-consuming and complex during the selection of positive lines, *Nicotiana benthamiana* was used as transient expression system. This has the advantage that the expression can be directly imaged within 2 days and the complicated selection process on RNA and protein basis can be avoided in the first instance. In addition, the perlucin-GFP construct was analyzed in tobacco to exclude organism specific inhibition of protein expression. Tobacco leaves were infiltrated as described in material and method section 2.2.4.6 and epidermal leaf cells were investigated after 48 h - 72 h (highest expression level). An overview about the different constructs is shown in Table 3.5.

Table 3.5: Leaves of *Nicotiana benthamiana* were infiltrated with *Argrobacterium tumefaciens* cultures containing the vectors mentioned below. The gene expression was investigated by fluorescence microscopical analysis. (*) The final vector construction was performed by student research assistant Anna Kreuder.

Plant Expression Vector
pMDC83-SP_{EXT3}-*perlucin*
pMDC83-SP_{PRP4}-*perlucin*
pMDC83-SP_{EXT3}-*ovocleidin-17* (*)
pMDC83-SP_{PRP4}-*ovocleidin-17* (*)
pMDC83-SP_{EXT3}-*N16N (*)*
pMDC83-SP_{PRP4}-*N16N (*)*
pALLIGATOR2 (GFP)

3.2.5.1 Microscopical Analysis of Plant Material

Confocal laser scanning microscopy (CLSM) and fluorescence microscopy was used to identify the protein expression by means of the fluorescence signal produced by OC17-GFP, N16N-GFP and perlucin-GFP. CLSM analyses revealed the expression and localization of the proteins OC17-GFP and N16N-GFP in the endomembrane system of

tobacco as show on the example of OC17-GFP (72h) in Figure 3.29. The fluorescent network filaments revealed the site of protein accumulation in the endoplasmic reticulum within the cell guided by a continious movement of fluorescent vesicles in the cell. Despite the obtained fluorescence signals, it was not possible to distinguish fluorescence signals derived from the cell wall area and from the inside of the cell. Subsequently, plasmolysis experiments were performed to overcome this limitation.

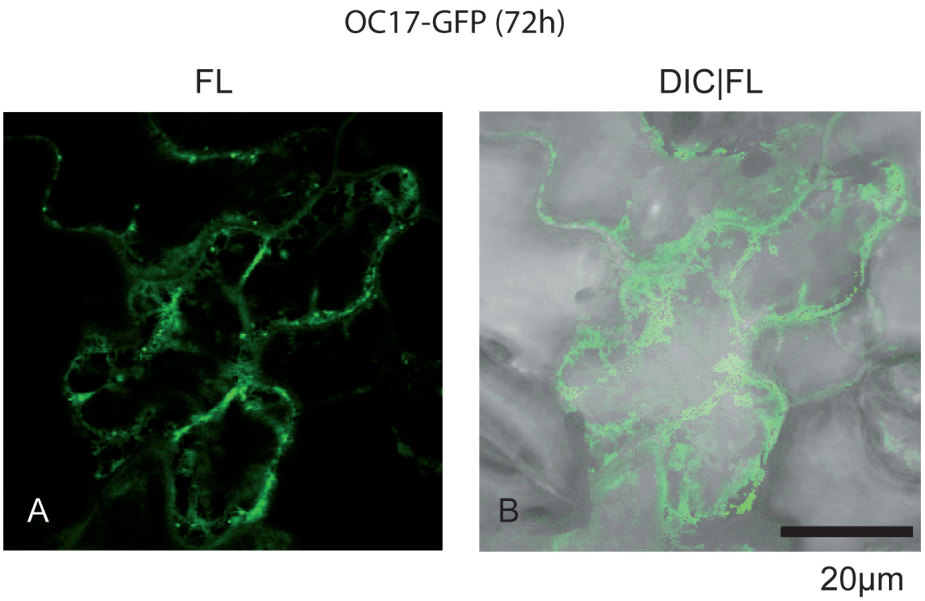


Figure 3.29: Expression of OC17-GFP was analyzed after 72 h in tobacco epidermal leaf cells. The fluorescent network and vesicles indicate a protein accumulation in the endoplasmic reticulum.

The plant material was analyzed before and after cell plasmolysis to investigate the protein expression pattern in the plant cell wall (material and method section 2.2.4.10). Figure 3.30 shows CLSM and light microscopical images before and after plasmolysis for direct comparison of the different protein expression pattern. Areas of particular importance are highlighted. PALLIGATOR2 served as transformation control due to its strong expression of a cytoplasmic GFP protein (Figure 3.30 A,F,K,O). It is noteworthy to mention that this construct contains a different promotor and cannot be directly compared to the constructs based on the pMDC83 system. It is expected that the expression pattern of cytoplasmic GFP after the plasmolysis experiment will be retained exclusively in the retracted protoplast. For comparison, the fluorescence signals derived from N16N-GFP and OC17-GFP are expected to be located in the endoplasmic compartments within the protoplast and in the cell wall area.

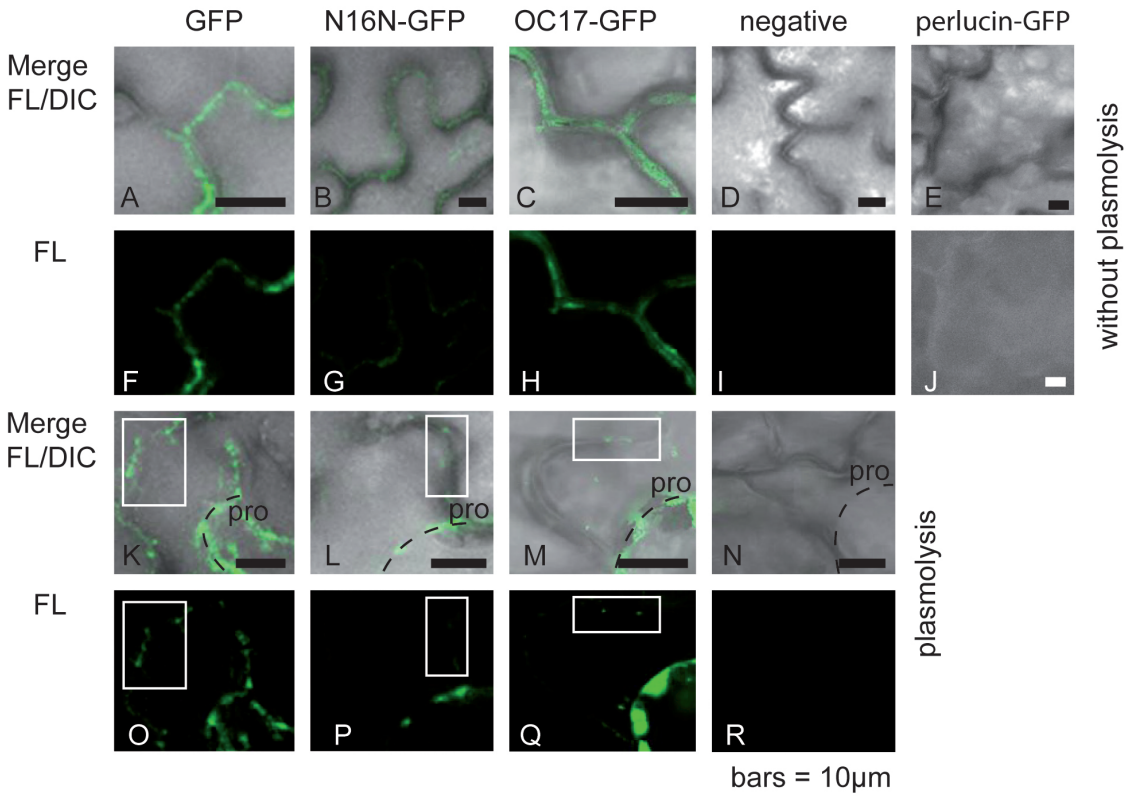


Figure 3.30: CLSM and fluorescence microscopy images of tobacco epidermal leaf cells (fluorescence mode (F-I, O-R) and overlay of DIC and fluorescence (A-D, K-N), transmission light image (E) and non confocal fluorescence image (J). Cells were displayed before (A-J) and after plasmolysis (K-R) expressing cytoplasmic GFP (A,F,K,O), endoplasmic N16N-GFP (B,G,L,P), endoplasmic OC17-GFP (C,H,M,Q), endoplasmic perlucin-GFP (E,J before plasmolysis) and the negative control (D,I,N,R) after 48 h. After plasmolysis and retraction of the protoplast (dashed lines, K-N) the strongest internal protoplast signals are visible in GFP-expressing cells (K,O). Additional signals remained from the cell wall area (highlighted). N16N-GFP and OC17-GFP showed strong signals in the retracted protoplast (L,P,M,Q) and less fluorescence signals remained from the cell wall area (L,P,M,Q, highlighted). Note that GFP served additionally as positive control for the transformation process. (pro) Protoplast, bar = 10 μm.

Tobacco epidermal cells are shown before plasmolysis (Figure 3.30 A-J) and after plasmolysis (K-R) expressing cytoplasmic GFP (A,F,K,O), endoplasmic N16N-GFP (B,G,L,P), endoplasmic OC17-GFP (C,H,M,Q), endoplasmic perlucin-GFP (E,J before plasmolysis) and the negative control (D,I,N,R). All adjustments were made such that cell wall autofluorescence in the negative control was kept at zero (D,I). Fluorescence images (F-I, O-R) and fluorescence and DIC overlay images (A-D and K-N) bore the remaining

fluorescence signals from the cell wall before plasmolysis (protoplast is appressed to the cell wall) and after the plasmolysis (protoplast is retracted from the wall). GFP (A,F) and OC17-GFP (C,H) revealed a strong fluorescence signal before plasmolysis on the cell borders, compared to N16N-GFP with a minor fluorescence signal (B,G). Under identical conditions no signal derived from the perlucin-GFP (E,J) sample.

After plasmolysis and retraction of the protoplast (Figure 3.30), dashed lines, K-N and O-R), the strongest signals from the protoplast were visible in the GFP-expressing cells (K,O) but also signals remained from the cell wall area (highlighted). N16N-GFP and OC17-GFP bore strong signals in the retracted protoplast (L,P,M,Q) and no signal remained from the negative control (N,R). Single fluorescence circles were observed in the cell wall area of N16N-GFP (L,P, highlighted) and OC17-GFP (M,Q, highlighted). In general, the selected DNA sequences encoding the signal peptides EXT3 and PRP4 had no influence on the signal intensity. In contrast, fluorescence signals remained from N16N-GFP expressing cells were less intense than from OC17-GFP.

3.2.5.2 Analysis of Plant Protein Extracts

The presence of GFP-fusion proteins was analyzed with an anti-GFP antibody in a Western blot procedure (Figure 3.31). Whole proteins were isolated from tobacco leaf cells containing OC17-GFP (46 kDa), N16N-GFP (34 kDa), GFP (29.7 kDa) or without foreign protein serving as negative control. The proteins were isolated as described in section 2.2.3.1 and separated on a SDS-PAGE. OC17-GFP and N16N-GFP analysis was performed for each signal peptide (EXT3 and PRP4), which will be removed by cellular processing from the precursor-protein during transfer to the endoplasmic reticulum due to specific proteases.

Both, the soluble and the insoluble protein fractions were investigated regarding the abundance of biomineralization proteins. The result of the soluble fraction is shown in Figure 3.31. The highlighted region indicates a weak band representing the full length variant of the OC17-GFP (46 kDa) on the correct size (upper arrow). In all samples, a strong protein band with an estimated molecular weight of 32 kDa was visible, which was also detected in the GFP-positive control. No specific signal of the N16N-GFP full length protein was detected (34 kDa, arrow).

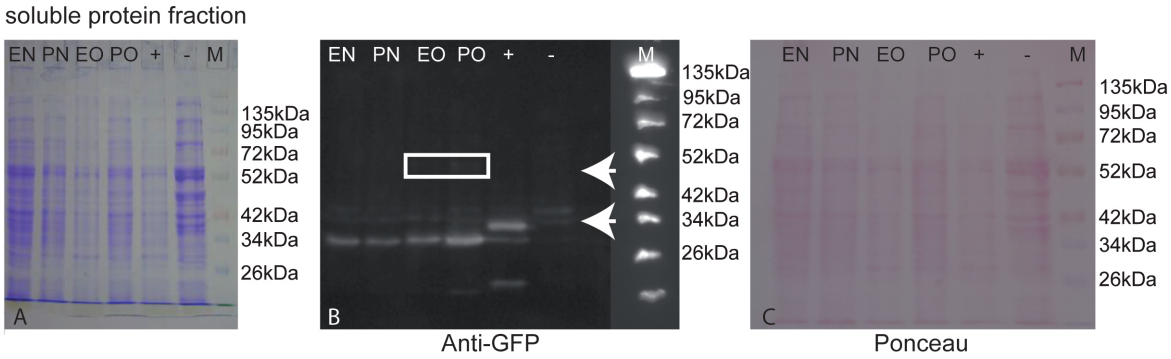


Figure 3.31: Whole protein extracts from transformed tobacco leaves (after 48h) were separated on a SDS-PAGE (A). OC17-GFP (46 kDa, arrow top) and N16N-GFP (34 kDa, arrow bottom) were detected with an anti-GFP antibody (B). The highlighted region indicates a weak band, which represented the full length variant of the OC17-GFP protein. A strong band was visible with less than 34 kDa in each sample. The GFP positive control (29.7 kDa) migrated on a size slightly smaller than 34 kDa. The successful transfer from the protein to the membrane was shown with a Ponceau staining (C). (EN) EXT3-N16N-GFP, (PN) PRP4-N16N-GFP, (EO) EXT3-OC17-GFP, (PO) PRP4-OC17-GFP, (+) GFP, (-) negative control, (M) molecular weight marker, Spectra^{TM} BR.

3.2.6 Summary: Expression of Biomineralization Proteins in Nicotiana benthamiana

Microscopical analysis bore a dominant signal for the OC17-GFP and N16N-GFP expression in the endomembrane system with similarities to the cytoplasmic GFP control protein. Protein analysis demonstrated an expression of N16N-GFP and OC17-GFP with a predominant GFP degradation product. Full length protein was detected in lower concentrations of OC17-GFP, but not for N16N. So far, no perlucin-GFP was detected in the tobacco plants. There was no obvious difference in the expression pattern observed for the usage of the different signal peptides PRP4 and EXT3.

4 Discussion and Outlook

4.1 Expression and Purification of Biomineralization Proteins in Bacteria

The synthesis of artificial biominerals and the biotechnological production of biomineralization proteins for fine-tuning the interaction with mineral phases is one of the most intruiging questions in biomineralization. In view of the fact that native and recombinant forms of biomineralization proteins always differ in their molecular structure and thus in terms of interactions with mineral interfaces, biotechnologically produced biomineralization proteins must be characterized carefully regarding their functionality.

4.1.1 Characterization of Recombinant GFP, GFP-Perlucin and Perlucin

The biomineralization protein perlucin from the soluble nacre fraction of mollusc *Haliotis laevigata* was expressed in a bacterial system (section 3.1.1 and 3.1.2) to study the protein activity *in vitro* in calcium carbonate precipitation assays (section 3.1.4.1). Three different constructs were designed for protein expression in the bacterial system: Perlucin, GFP-perlucin and GFP. All gene sequences were fused to an affinity HIS_6 tag sequence, for facilitated protein purification. The results revealed that GFP and GFP-perlucin were obtained under native conditions in a one step purification procedure. The amount of recombinant protein was increased due to the optimization of the bacterial growth conditions, whereas the native purification of perlucin without GFP was not achieved either under standard or optimized conditions.

Two major fragments were observed in the one step eluate of GFP-perlucin and mass spectrometric analysis revealed that several variants of the GFP-perlucin were obtained including two truncated variants (29.8 kDa - 30.3 kDa) and the full length variant with the two terminal repeat sequences (46.8 kDa, section 3.1.3.1). SDS-PAGE analysis

revealed that the amount of purified truncated variants was higher than the full length variant (Figure 3.3). The formation of truncated protein variants was promoted due to the lacking of codon usage optimization, which was here optimized for the expression in the plant organism.

The growth profile of perlucin, GFP-perlucin and GFP expressing cultures showed that the expression of the pure perlucin lead to reduced growth rates, indicating a toxic effect, which is not reported for the native extracted perlucin from *Haliotis laevigata* [Weiss2000] (Figure 3.4). It cannot be excluded that the growth reduction of perlucin expressing cells is influenced by the lectin properties of perlucin, which may interfere with cell compounds or processes in the bacterial organism such as cell division and cell wall assembly. The fusion to the GFP-protein allowed to overcome this effect and enhanced the protein solubility. The recombinant form of perlucin misses the post-translational glycosylation of an asparagine residue, which was detected in the native perlucin [Weiss2000], [Mann2000]. One disadvantage of the bacterial expression system is the limitation regarding posttranslational modifications (PTM) and it seems most likely that the missing glycosylation might impair the protein structure and function in several ways, which will be discussed later. However, the insoluble perlucin was not used further.

The GFP-perlucin variants were separated with native PAGE (section 3.1.3.2) and size exclusion chromatography (section 3.1.3.3) for subsequent upscaling. In addition, both methods allowed an independent investigation of the protein agglomeration behaviour. It was shown that truncated variants can be separated from the full length variant forming small protein agglomerates of few individual polypeptide chains. The full length variant in contrast built a multi-protein complex with the truncated variant as demonstrated by size exclusion chromatography and SDS-PAGE analysis (Figure 3.8). The terminal repeat sequences of perlucin are suggested to have a strong influence on the agglomeration behaviour, which is assured through this study. Further agglomeration analysis was performed in dependence on their first ionic interaction partner (HCO_3^- or Ca^{2+}) with the field-flow fractionation system. HCO_3^- - protein interaction induced lower protein agglomerates, than Ca^{2+} (section 3.1.4.3). This may be explained (a) by the higher molecular weight of GFP-perlucin resulting in a larger hydrodynamic diameter and (b) a higher sensitivity of GFP-perlucin to changes in the ionic environment. Further (c) disulfide bonds (perlucin full length variant contains 6 cysteine residues) were suggested to promote the complex formation under native conditions. To evaluate the role of disulfide bonds, the fractionation analysis will be performed in the presence

of dithiothreitol (DTT), a strong reducing agent for disulfide bonds. The agglomeration of biomineralization proteins is often part of their activity and therefore GFP and GFP-perlucin one step eluates as well as the single protein variants were used for protein activity assays to investigate their influence on calcium carbonate precipitation.

4.1.2 Influence of Recombinant GFP and GFP-Perlucin on Calcium Carbonate Formation

The activity of GFP and GFP-perlucin one step eluate and a lectin control protein was analyzed in a simple calcium carbonate supersaturation bulk assay (section 3.1.4.1). Two precursor solutions ($CaCl_2$ and $NaHCO_3$) were combined and the assays differed in the precursor solution to which the protein eluate was added first (protein-Ca^{2+} and protein-HCO_3^--precursor). It was shown that GFP as well as GFP-perlucin one step eluate has an inhibitory influence on calcium carbonate formation at high concentrations (exceeding 2 μg/ml *in vitro*). The effect was more distinct if HCO_3^- was applied as first ionic interaction partner (Figure 3.9). The inhibitory effect was not described in literature so far but seems to arise from the GFP domain rather than from perlucin (Table 3.1).

Precipitates with characteristic morphologies were observed in all samples and the question how the proteins are interacting with the mineral phase was analyzed using fluorescence microscopy analysis due to the internal fluorescence of GFP. Spherical and rhombohedral forms were obtained if Ca^{2+} was used as first ionic interaction partner (Figure 3.10). Flower shaped calcium carbonate crystals (Figure 3.11) were grown at high protein concentrations with an accumulation of GFP and GFP-perlucin at the interfaces between crystalline building blocks (Figure 4.1).

The size of the macromolecules prevented their incoperation into the ion lattice of the crystals but allowed them to intercalate with the crystal surfaces (Figure 4.1). This lead to modified physical properties and different mineral shapes [Mann2001]. The inhibitory effect is known for the addition of organic additives such as acidic polysaccharides or proteins in the range of 1 mg/ml [Mann2001]. Raman analysis revealed the presence of calcite, a stable form of calcium carbonate. No stabilization of aragonite or vaterite was observed, which was also previously described for perlucin [Blank2003].

To follow the question why the inhibitory effect of GFP is more intense if protein first interacts with HCO_3^-, the local ion dispersion has to be considered. According to former analysis, the growth of the mineral phase depends on the ion excess adsorbed on the surface of the crystal nuclei [Lippmann1973].

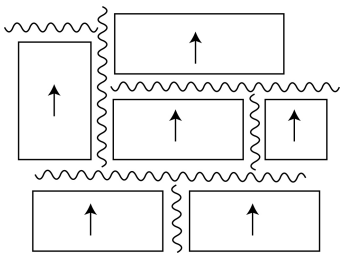


Figure 4.1: Schematic illustration of an inorganic mineral that intercalates with organic molecules at the interfaces. This mechanism is suggested for GFP- and GFP-perlucin-mineral interaction at high protein concentrations if HCO_3^- acts as first ionic interaction partner (scheme has been modified from [Mann2001]).

To aim a better understanding, the calcium binding sites in GFP were predicted with *in silico* analysis (section 3.1.4.2). It was predicted that GFP is able to bind calcium ions on the loop I and II region of the barrel structure, which are rich in acidic amino acids (appendix Figure 5.5). The accumulation of positively charged calcium ions in the loop regions indicates an influence on the solubility equilibrium in the supersaturated calcium carbonate solution, but it may also promote the formation of crystal nuclei on the surface of the GFP-protein. To verify the hypothesis, it has to be evaluated, whether GFP shows an *in vitro* calcium binding capacity. In addition, the *in vitro* calcium binding property of recombinant GFP-perlucin has to be analyzed to exclude the deficiency of the calcium binding property as previously observed after expression of the pearlin protein from the oyster *Pinctada margaritifera* in a bacterial system [Montagnani2011].

Furthermore, the question arises, whether the GFP and the GFP-perlucin variants are involved in stabilization of crystalline precursor phases. The formation and stabilization of amorphous calcium carbonate is a common strategy in biomineralization events [Addadi2003] and the influence of biomineralization proteins on the formation of prenucleation species is under recent discussion [Gebauer2008]. Therefore, the single protein species will be anlyzed in cooperation with the group of Prof. Helmut Cölfen (University of Konstanz), which allows to investigate a protein dependent appearance of mineral precursor species [Gebauer2008].

Recombinant GFP and perlucin are suggested to be calcium carbonate "active" in a bivalent role such that GFP acts as inhibitor and perlucin as promotor of calcium carbonate formation. It seems that GFP is the dominant domain in the GFP-perlucin fusion protein. A reduced promoting effect is suggested to be decreased through the excess of truncated protein variant in the one step protein eluate. The calcium carbonate precipitation assay of single GFP-perlucin variants separated by size exclusion

chromatography will be analyzed in future experiments.

The influence of GFP on calcium carbonate formation was not described so far, however, native isolated perlucin from *Haliotis laevigata* is known to promote calcium carbonate formation [Weiss2000], [Blank2003]. Note that the recombinant perlucin missses a glycosylation compared to the native isolated form. To summarize, the missing glycosylation might impair the perlucin protein in two different ways:

- Firstly, perlucin has a reduced solubility compared to the native protein and the fusion to GFP is able to mimick these glycosylation to some extent and allows the native purification.
- Secondly, PTMs influence degradation processes and can protect the protein against proteinase digestion [Vlasak2009]. The lacking glycosylation may promote the cleavage of the protein as observed here.

4.1.3 Characterization of Metastable GFP and GFP-Perlucin Vaterite Composites

A second assay was performed to study the interaction of GFP and GFP-perlucin with the mineral phase regarding the influence of reaction volume (section 3.1.4.4). After 2 h of incubation, spherical particles coexist in controls without protein and in assays with low concentration of protein together with typical rhombohedral and octahedral crystals [Weber2012]. Raman spectroscopy revealed that the spherical particles consisted of vaterite (Figure 3.16), whereby all other forms consists of calcite. The vaterite particles had morphological similarity with those precipitated *in vitro* using comparable concentrations of natively purified Chinese soft-shelled turtle (*Pelodiscus sinensis*) eggshell protein, pelovaterin [Lakshminarayanan2005] and similar microparticles precipitated in the presence of chemical fluorophores [Pai2011]. Comparable morphologies and size ranges were obtained by using peptoids, but yielded calcite crystals [Chen2011].

The vaterite particles disappeared after 24-72 h (Figure 3.17), which were only fluorescent when the precipitation was performed in the presence of GFP [Weber2012]. Furthermore, vaterite might transform into aragonite and calcite under ambient conditions [Beck2010], [Lippmann1973].

The favourable incorporation of Ca^{2+} loaded GFP and GFP-perlucin into vaterite does not mean that GFP or GFP-perlucin actually drives vaterite formation. This may indicate a combination of two secondary effects: The depletion of Ca^{2+}, which leads to an excess of HCO_3^- and the favoured incorporation of Ca^{2+}/GFP on CO_3^{2-} terminated

crystal faces with less regular structure, which leads to vaterite [Weber2012]. This may result in a slightly different crystal growth rate and 3D geometry such as shown for the protein-vaterite spheres, which revealed a reduced particle area compared to the control. The temporary interaction of GFP domains with vaterite is indeed important for biomimetic mineral precursor strategies, which may be applied in order to control mineral morphologies. Due to the metastability of vaterite, insoluble protein skeletons with characteristic morphologies and optical properties remained attached to the slide. There was no difference between GFP and GFP-perlucin truncated observed in this assay, indicating a similar protein-mineral bevahiour, whereas the full-length variant of GFP-perlucin had the tendency to agglomerate, which was visible as fluorescent precipitates in the background of the microscopic slide. Compared to small grained precipitates in the sample with high protein concentrations of GFP and GFP-perlucin, the full length precipitates had a sticky character and was not removed from the surface of the microscopic slide with compressed air. In addition, slightly fluorescent vaterite spehres were observed in the same wells (Figure 3.15). The complex formation of GFP-perlucin truncated with the full length variant indicates an increase in solubility of the protein complex. During protein interaction with mineral phase and the potential incoperation of GFP-perlucin truncated into the vaterite particles, the full length variant may have stronger tendency to precipitate. To guide a protein to its site of activity in the mollusc shell or consequent degradation of a precursor species to a functional protein [Dodenhof2010], is suggested to be a strategy how perlucin acts under natural conditions.

4.1.4 Summary and Outlook

This observations let suggest that a soluble protein such as GFP which is not evolutionarily optimized with respect to mineralization and mechanical function, but rather for optical purposes, is able to mimick certain aspects of natural biomineralization proteins. Synthetic biominerals from large volume assay seem unlikely to form a mixed crystals [Kahr2001] and it would be worth checking whether or not the crystal lattice structure varies systematically with the composition of organic in both assays, which will be achieved with x-ray based analysis. The mechanical properties will be further studied with nanoindentation analysis to investigate a change in stiffness and strength in the artificial biominerals.

It was also demonstrated that the influence of reaction volume has a major impact on crystal formation. In addition, the 30 μl assay can be used as fast screening system

for different protein variants and concentrations and the advantage compared to the 12 ml assay (large volume assay) includes that less protein has to be prepared for protein activity analysis. To adress the question, if the protein is a promotor or an inhibitor of calcium carbonate formation, the 12 ml assay has to be used.

In further approaches, the additives will also be used separately for incorporation into nanoparticles as shown for GFP-silica particles, which can be used for microscopical application such as the study of nanoparticle-cell interaction.

Delayed precipitation and metastable inorganic mineral phases are of general research interest. Fine-tuning the properties of such proteins with respect to the metastability of mineral precursors would then be a promising, versatile route to obtain multifunctional organic coatings and composite materials. The features of GFP-perlucin as revealed in part I of the thesis by different *in vitro* assays would be interestingly to investigate in plant cell walls as kind of *in vivo* assay.

4.2 Expression of Biomineralization Proteins in Plants

The heterologous expression of proteins in the plant organism opens a versatile route towards a broad spectra of applications such as the production of pharmaceuticals, antibodies or industrial proteins (section 1.5.1). Due to the advantages of plants as bioreactors for protein expression, the question was adressed, whether a selection of biomineralization proteins can be successfully synthesized by plant organisms. The cell wall was used as target compartment, because of its important functions in mechanically supporting the plant cell and mimicking the natural biomineralization matrix in certain aspects.

The cell wall and the organic matrix of mollusc shells are deposited extracellularly and serve as major structural polymer (chitin in molluscs, cellulose in plants). The assembly of the organic framework is highly controlled, which can be active or passive via self-aggregation. The cell wall as well as the organic matrix forms a hydrogel (silk proteins in molluscs, pectins in plants) and both organisms achieve a change in strength, due to modification of protein composition, which can became insoluble after delivery. It is noteworthy to mention that these proteins often obtain repetitive sequences. Also the nacre protein perlucin carries repetitive sequences. The cell wall and the mollusc shell contain calcium, which can interact with acidic molecule species giving structural support. In plants the pectin network is mainly responsible for reversible calcium-protein interaction serving as important mechanism during cell expansion. A large

variety of modified proteins and polysaccharides were observed to enable the synthesis of complex networks and offer a potential binding site for biomineralization proteins with a lectin-binding domain such as known for perlucin. *Arabidopsis thaliana* and *Nicotiana benthamiana* plants were analyzed regarding their ability to express biomineralization proteins.

4.2.1 Expression of Perlucin-GFP in Arabidopsis thaliana

The first organism, which was selected for heterologous protein expression was *Arabidopsis thaliana* with the aim to generate perlucin-GFP transgenic lines. Perlucin seemed to be the most promising candidate due to its lectin-binding capacity, which enables them to bind cell wall polysaccharides (section 1.1.4.1). To achieve the expression, three different signal peptides were selected for the expression in the plant cell wall (EXT3, PRP4 and chitin-hydrolase). The sequences originated from cell wall proteins of *Arabidopsis thaliana*, which were previously characterized and associated to three different families of cell wall proteins (section 3.2.1).

The investigation of the gene- and protein expression level in the t2 generation showed that neither RNA nor protein was detected with standard molecular biological methods. However, a successful expression cannot be excluded assuming an expression rate under the detection limit (section 3.2.3.3). In addition, the t2 phenotype investigations revealed a reduced germination rate in some identified lines compared to wild type plants (section 3.2.3.3).

Subsequently, the seed morphology was analyzed with scanning electron microscopy, which revealed morphological changes including the deformation of the whole seed as well as surface alterations (Figure 3.25). Interestingly, ESEM images of seed coat epidermis of a dry and hydrated seed surfaces of knouck out mutants *df1-1* and *df1-2* bore different surface morphologies. Both genes are identified to play a role in the mucilage synthesis in *Arabidopsis* seeds [Vasilevski2012]. The mucilage composition analysis revealed pectins as the major compound [Macquet2007]. The interaction of calcium binding proteins in the plant organism is suggested to interfere with the assembly of the pectin network during seed formation. Some lines owning a reduced germination rate were correlated with lines bearing the deformed seeds.

In addition, the mean root length of t2 plants were measured and three lines were identified, which had either reduced or elongated mean root length compared to wild type (section 3.24). The t3 generation revealed a reduced germination rate as well as a

reduction in the seed formation rate in siliques of two lines, which were different from the identified lines in the t2 generation (Figure 3.28). However, further phenotype analysis did not bear differences compare to the wild type. The measurement of calcium content in the seeds of wild type and mutants would be a possible strategy to give an indirect proof that calcium homeostasis is influenced by foreign gene expression.

Since the mollusc shell consists of 95 % inorganic calcium carbonate and only of 5 % organic, it has to be assumed that the impact of organic molecules is highly efficient and minimal amounts of perlucin-GFP may have an influence on plant growth and development. In particular, if the perlucin-GFP expression interferes with critical key steps in the plant organisms such as cell wall formation, lethal effects are suggested to occur. This assumption is assured through the phenotypical observations as described above. It seems that *Arabidopsis* plants react on the expression of perlucin-GFP with morphological changes, mainly regarding seed development. The seed formation and plant growth is a very fine-tuned process in the plant, especially at very early stages [Bonza2000]. The ability of perlucin not only to bind calcium ions and sugar residues, but also a reduced solubility has to be considered. The supposed toxic influence of perlucin with respect to the growth of bacteria as described above is an additional indication for a strong influence on cell developmental processes. Although there is no proved explanation for these findings, different hypotheses will be discussed in the following section.

4.2.2 Possible Effects of Perlucin-GFP Expression on Plant Development

4.2.2.1 Secretory Pathway and Calcium

The plant secretory system is a highly complex network and the main function includes the synthesis and transport of glycoproteins, proteoglycans, lipids and branched polysaccharides for the cell wall and cell plate formation (section 1.5.3) [Albersheim2011]. The endoplasmic reticulum (ER) as part of the endomembrane system is distributed over a cell with different functional areas. The molecule processing is continued in the Golgi apparatus and the *trans* Golgi network (TGN), which completes the biosynthesis of polysaccharides and their delivery to the cell surface or vacuole via vesicular transport.

In addition, the distinct control of calcium homeostasis has a high impact on plant physiology such as photomorphogenesis, stomatal closure, root hair growth and pollen

tube development [Sanders1999]. Calcium plays also a major role in the secretory pathway such as delivery of vesicles to the cell surface, attachment of vesicles to the plasma membrane and the formation of fusion pores. Calcium acts as second messenger and it was suggested that waves and frequencies of cytosolic Ca^{2+} content define the signals. To accomplish the Ca^{2+} pulsion, channels in the plasma membrane, the endomembrane system and the vacuole are responsible to maintain the specific concentration (Figure 4.2 [Sanders1999]).

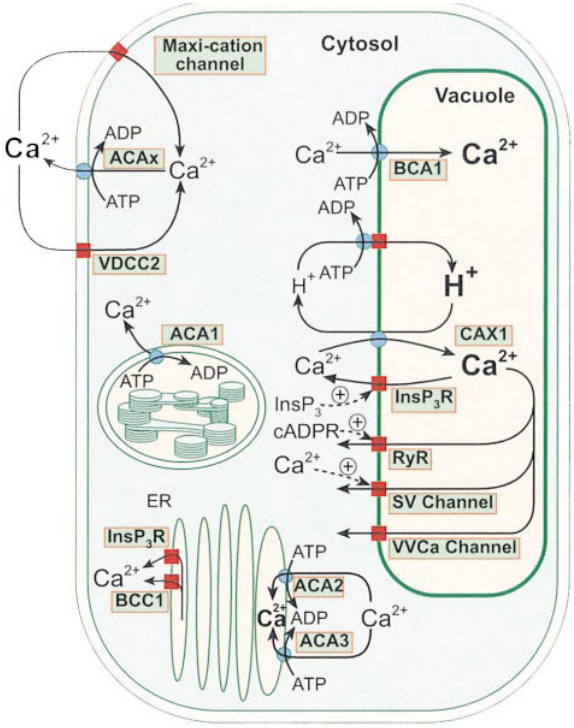


Figure 4.2: Overview of major calcium transport pathways in the plant organism [Sanders1999]. (ACA) autoinhibited calcium dependent ATPase, (CAX) Ca^{2+}/ H^+ antiporter, (BCC1) *Brionica* Ca^{2+} channel, ($InsP_3R$) putative $InsP_3$ receptor, (RyR) putative ryanodine receptor, activated by cADPR, (SV channel) slowly activating vacuolar channel, (VDCC2) voltage-dependent Ca^{2+} channel, (VVCA) vacuolar voltage gated channel, (red squares) Ca^{2+} permeable ion channels, (blue circles) energy dependent transport.

Proton pumps (V-ATPase, V-PPase) are important to control the pH in the Golgi compartments and defects in the catalytic unit of a *Arabidopsis* V-ATPase activity results in male and partial female gametophytes lethality. A shift in the ion equilibrium may influence the turgor pressure negatively, which is important for vesicle secretion and recycling of cell wall compounds [Albersheim2011].

Fourteen calcium pumps have been identified in *Arabidopsis* genome with 10 encoding Ca^{2+} ATPases, which have a calmodulin-dependent autoinhibitory domain on the N-terminus (ACA) [Bonza2000], [Baxter2003]. The effect of knock out mutants of the plasma membrane ATPase ACA9 on the pollen tube development was demonstrated with a partial sterility revealed by a non-mendelian segregation of homozygous plants [Bonza2000]. The plants had shorter siliques and 80 % less seeds in the siliques, which was correlated with growth defects of the pollen tubes [Bonza2000]. In addition, analysis of a knock out mutant of the *ACA10* gene, which encodes a further plasma membrane Ca^{2+} ATPase, demonstrated its unique function in regulation of plant elongation and the specification of key aspects in the development of inflorescence structures [George2008]. The modified distribution pattern of calcium and other elements within seeds by perturbed CAX activity was investigated in *cax1* and *cax3* loss-of-function lines using synchrotron x-ray fluorescence microscopy [Punshon2012]. A crucial role of nutrient localization for normal growth and development was demonstrated previously [Kim2006].

Although a minor statistical analysis regarding seed formation was obtained for the perlucin-GFP plants, the observations indicates an interference with the calcium transport system in the plant.

Regarding cell wall compounds, calcium plays a major role in the formation of the pectin network and cell expansion during primary cell wall formation. Preferably, the middle lamella is connected to the cell wall of two neighbouring cells via calcium bridges if the acidic homogalacturonanas (HGA) are non-methylesterified (introduction 1.4.3) [Albersheim2011]. The reaction is processed by pectin methylesterase enzymes (PME) and a consequent overexpression of a PME inhibitor achieved a less content of non-methylesterified HGA's. *Arabidopsis* plants overexpressing the PME inhibitor showed an enhanced growth, but also a less structured hypocotyl texture [Lionetti2010].

The deposition of callose (linear β-linked glucose polymer) during cell growth and development may also be negatively influenced due to the overexpression of perlucin. Callose is directly synthesized by plasma membrane enzymes and the synthesis can be induced by calcium. For example, callose formation is required for a correct pollen development and male sterile pollen can be obtained due to an inhibition in callose deposition [Worrall1992]. The interference of a lectin protein with various calcium-regulated processes is suggested to have a high impact on primary cell wall formation and to affect the cell growth at very early stages.

4.2.2.2 Lectins

Lectins can be linked to a large number of different biological processes over all kingdoms of life such as from viruses and bacteria up to plants and animals [Loris2002]. In plants, their functions cover protection from pathogens, establishment of symbiosis, storage, growth regulation and the adaption to environmental conditions [Gabius2004]. The transfer of biological information is regulated by protein-carbohydrate interactions.

The advantage of carbohydrates as signalling compound is given by their high complexity including branching of functionalized groups and their flexibility. The signal transduction occurs due to the binding of sugar residues to lectin proteins. Lectins belong to a diverse family of proteins such as known for the calcium dependent family, the C-type lectins [Loris2002]. The loss of the metal ion binding induces a conformational change and the functionality of the carbohydrate binding site is destroyed. The mechanism depends on a *cis-trans* isomerization of peptide bonds. It is postulated that these slow processes facilitate the sorting of carbohydrate-bounded ligands in the endosome [Ng1998]. However, the diversity of lectins include differences between plant and animal lectins, such that the binding types of carbohydrate and metal ion is rather direct in animals as non direct in legumes [Loris2002].

The lectin-carbohydrate signalling is tightly controlled for internal genes. However, a foreign overexpressed lectin protein has the ability to disturb cellular processes such as protein translocation to the endomembrane system or protein sorting due to an undesired interaction with internal sugar residues.

4.2.2.3 Protein Degradation-Cellular Responses

The regulation of the cellular processes has to be controlled to ensure correct processing of the constituents and to avoid the delivery of misfolded proteins to the cell surface, which may affect cell-cell communication [Roemisch2005]. One of these systems is the ER associated degradation system (ERAD) already mentioned in the introduction. Terminally misfolded proteins were transported across the ER membrane into the cytosol, where they can be degraded via proteasomes [Roemisch2005]. A second system, which is connected to ERAD is the unfolded protein response (UPR). If the ERAD system is not functional, the UPR is induced and permanently active. In contrast the UPR system can promote the ERAD system [Bukau2006] and the loss of both systems lead to a decrease in cell viability.

The activated UPR system senses misfolded molecules via transmembrane proteins and can induce a decrease not only in protein translation but also an increase in protein synthesis, which facilitates a correct folding of the proteins, such as chaparones. It is known that the sugar residues attached to proteins serve as signal for the state of protein folding [Bukau2006]. However, the UPR system is also able to induce cell apoptosis [Bukau2006]. The overexpression of proteins in transgenic plants with protein accumulation and protein misfolding is suggested to induce ERAD as well as the UPR system and initiate a consequent degradation of proteins. Due to a feedback mechanism, RNA and protein degradation processes may be promoted, which would explain the missing proof for the abundance of RNA and protein in the perlucin-GFP plants. It would be worth to investigate the induction of the ERAD and UPR system in following projects by gene expression analysis of UPR/ERAD marker genes.

4.2.3 Expression of Perlucin-GFP, N16N-GFP and OC17-GFP in Nicotiana benthamiana

Regarding the findings of the expression of perlucin-GFP in *Arabidopsis*, N16N-GFP, OC17-GFP and perlucin-GFP were transient expressed in epidermal leaf cells in *Nicotiana benthamiana*. The protein expression was investigated using fluorescence microscopy and Western blot analysis of whole cell extracts. Transgenic *Arabidopsis* plants usually produces the protein in all developmental stages, whereas epidermal leaf cells of *Nicotiana benthamiana* are transformed in 3-4 leaf stage and the cells are fully developed such that in general different conditions are provided. Similar to the obtained results of perlucin-GFP in *Arabidopsis*, no protein was detected in *Nicotiana benthamiana* so far. Therefore, the influence of perlucin seems to be independent of the plant organism, indicating fundamental problems in protein biosynthesis as discussed before.

Fluorescent signals from N16N-GFP and OC17-GFP expressing plants were detected in the endomembrane system of epidermal leaf cells (Figure 3.29). This indicates that the signal peptides for protein targeting to the plant cell wall are functional regarding protein delivery to the secretory pathway. To prove this hypothesis, co-expression analyses with Golgi marker proteins must be perfomed to study their expression pattern in the leaf cell.

The Western blot analysis demonstrated the abundance of a dominant degradation product and minor amounts of full length variant. The degradation product of GFP is suggested to overlay with the full length variant of N16N. These findings indicate that the

major part of the protein is separated in at least two fragments and only the GFP fragment is detected with the here used anti GFP antibody in the Western blot procedure (Figure 3.31). Due to the abundance of the identical degradation product in the GFP positive control, it can be suggested that the sequence contains a natural protease recognition site. The usage of gradient gels would be useful to obtain a better separation for the N16N with a molecular weight of 34 kDa. Due to the C-terminal fusion of GFP to the biomineralization proteins, their production is most likely. Surprisingly, the degradation process is more distinct for N16N-GFP and OC17-GFP than for the positive control. This indicates an increase in proteolytic activity, which is suggested to be an indirect confirmation that the plant organism reacts specific on different proteins with an adaption of cellular processes. Different buffer systems containing protease inhibitors and less mechanical stress during protein isolation was not able to reduce the protein degradation.

In addition, the signals derived from OC17-GFP transformed leaf cells were stronger compared to N16N-GFP. This fact is also an indication for a protein specific reaction, which affects the protein expression level although an identical 2xCaMV35s promotor and cloning strategy was used for all constructs (section 3.2.2). Therefore, a down-regulation of the protein is suggested for perlucin-GFP via previously described mechanisms (ERAD and UPR). Nevertheless, a successful transfer of the proteins to the cell wall cannot be excluded.

To analyze the expression pattern in more detail, plasmolysis experiments were executed and revealed similar fluorescence pattern for N16N-GFP, OC17-GFP and the control protein (cytosolic GFP), with fluorescent circles in the cell wall region after the plasmolysis (Figure 3.30). This indicates that the signals remain from cytoplasmic strands containing ER domains, which connect two neighbouring cells via plasmodesmata. Furthermore, it can be speculated that the plasmolysis experiments itself have a negative influence on the protein identification. The plasmolysis may promote a protein leakage out of the cell wall and small amounts of the full length proteins would not be detectable. This effect is suggested to play a major role because both proteins do not contain a functional lectin-binding domain (OC17 has a domain with similarities to a C-type lectin). For comparison, a stable fluorescence pattern was obtained after plasmolysis for the heterologous expressed plasmodesmata-associated glycoproteins RGP2 [Sagi2005]. Further, the acidic pH of the cell wall may also reduce the fluorescence signal from GFP [Ellis2008] or the GFP itself would disturb a correct delivery of the proteins, which was described in case of a C-terminal fusion protein to the plasma membrane

protein CESA1 [Tian2004].

4.2.4 Summary and Outlook

Perlucin, N16N and OC17 are different in amino acid composition, physical properties, structure and function although they are linked to calcium carbonate biomineralization processes. To facilitate the protein biosynthesis, the codon usage of all gene sequences was optimized for the plant organisms. Compared to OC17, perlucin classified as functional C-type lectin, seems to interfere most likely with cellular processes and N16N can modify its structure in dependence on calcium. Further, N16N has the tendency to form oligomers if the pH is changed, which is suggested to be responsible for a minor expression level compared to OC17. Nevertheless, the expression of biomineralization proteins was achieved the first time in a plant organism and with the knowledge gained in this study, the strategy for the expression of biomineralization proteins can be optimized using the following options:

- Firstly, referring the cloning strategy, the 2xCaMV35s promotor has to be changed against an inducible promotor. This would help to overcome possible embryonic lethal effects, which are induced by the foreign gene expression of a potential toxic protein. Here, the comparison between an expression in seeds and after seed germination would be interesting to investigate regarding a toxic effect of the protein.
- Secondly, the protein fusion to the GFP domain has to be changed against a different tag to overcome the proteolytic degradation and the pH sensitivity, which is critical in the case of a very low expression level [Ellis2008]. It is known that YFP is more resistant against pH variations and would therefore be a potential candidate [Tian2004].
- Thirdly, to overcome leaching effects in the cell wall after protein delivery, the proteins can be fused to a cellulose binding protein, which would be very efficient in binding of extracellular cellulose fibres.
- Fourthly, with the gained knowledge of the *in vitro* studies of recombinant produced GFP and GFP-perlucin, an interference of the GFP-domain with the calcium homeostasis can also not be excluded. The usage of a different protein-tag system (such as hemagglutinin-tag) would allow to investigate the influence of the GFP domain.

- Fifthly, the usage of a hairy root system combined with an inducible promotor for foreign protein expression would allow to analyze strong proliferating cells which were easy to observe with microscopical techniques.

Real-time PCR analysis of marker genes, which correspond to different cellular pathways or a microarray analysis would show, which process is affected by the expression of biomineralization proteins. Furthermore, the application of different chemicals, which inhibit specific cellular processes, will also help to understand the findings of this study and identify affected plant cellular pathways. The expression of biomineralization proteins in plants is suggested to be improved due to the usage of plant mutants as host system such as plants with higher calcium content in the cell wall. This study served as first proof-of-principle and provided a number of guidelines for future approaches to achieve long term goals regarding a synthesis of new and sustainable materials such as the structural modification of the plant cell wall.

Bibliography

[Addadi1985] L. Addadi and S. Weiner. Interactions between acidic proteins and crystals: Stereochemical requirements in biomineralization. *Proceedings of the National Academy of Sciences*, 82(12):4110 – 4114, 1985.

[Addadi2003] L. Addadi, S. Raz, and S. Weiner. Taking advantage of disorder: Amorphous calcium carbonate and its roles in biomineralization. *Advanced Materials*, 15(12):959 – 970, 2003.

[Addadi2006] L. Addadi, D. Joester, F. Nudelman, and S. Weiner. Mollusk shell formation: A source of new concepts for understanding biomineralization processes. *Chemistry - A European Journal*, 12(4):980 – 987, 2006.

[Albersheim2011] P. Albersheim, A. Darvill, K. Roberts, R. Sederoff, and A. Staehelin. *Plant Cell Walls*. Garland Science, Taylor & Francis Group, LLC, pages 86 - 91, 311 - 313, 2011.

[Arias1991] J. L. Arias, M. S. Fernandez, J. E. Dennis, and A. I. Caplan. Collagens of the chicken eggshell membranes. *Connective Tissue Research*, 26(1 - 2):37 – 45, 1991.

[Arias1993] J. L. Arias, D. J. Fink, S.-Q. Xiao, A. H. Heuer, and A. I. Caplan. Biomineralization and eggshells: Cell-mediated acellular compartments of mineralized extracellular matrix. volume 145 of *International Review of Cytology*, pages 217 – 250. Academic Press, 1993.

[Arioli1998] T. Arioli, L. Peng, A. S. Betzner, J. Burn, W. Wittke, W. Herth, C. Camilleri, H. Höfte, J. Plazinski, R. Birch, A. Cork, J. Glover, J. Redmond, and R. E. Williamson. Molecular analysis of cellulose biosynthesis in arabidopsis. *Science*, 279(5351):717 – 720, 1998.

[Arnott1970] H. J. Arnott and F. G. E. Pautard. *Biological calcification: Cellular and molecular aspects*. Appleton-Century-Crofts, 1970.

[Bacic1988] A. Bacic, P. J. Harris, and B. A. Stone. *The biochemistry of plants*, pages 297 – 371. Academic Press, New York, USA, 1988.

[Bakker2001] H. Bakker, M. Bardor, J. W. Molthoff, V. Gomord, I. Elbers, L. H. Stevens, W. Jordi, A. Lommen, L. Faye, P. Lerouge, and D. Bosch. Galactose-extended glycans of antibodies produced by transgenic plants. *Proceedings of the National Academy of Sciences*, 98(5):2899 – 2904, 2001.

[Baxter2003] I. Baxter, J. Tchieu, M. R. Sussman, M. Boutry, M. G. Palmgren, M. Gribskov, J. F. Harper, and K. B. Axelsen. Genomic comparison of p-type atpase ion pumps in arabidopsis and rice. *Plant Physiology*, 132(2):618 – 628, 2003.

[Beck2010] R. Beck and J.-P. Andreassen. The onset of spherulitic growth in crystallization of calcium carbonate. *Journal of Crystal Growth*, 312(15):2226 – 2238, 2010.

[Belcher1996] A. M. Belcher, X. H. Wu, R. J. Christensen, P. K. Hansma, G. D. Stucky, and D. E. Morse. Control of crystal phase switching and orientation by soluble mollusc-shell proteins. *Nature*, 381(6577):56 – 58, 1996.

[Bensmihen2004] S. Bensmihen, A. To, G. Lambert, T. Kroj, J. Giraudat, and F. Parcy. Analysis of an activated abi5 allele using a new selection method for transgenic arabidopsis seeds. *FEBS Letters*, 561(1–3):127 – 131, 2004.

[Blank2003] S. Blank, M. Arnoldi, S. Khoshnavaz, L. Treccani, M. Kuntz, K. Mann, G. Grathwohl, and M. Fritz. The nacre protein perlucin nucleates growth of calcium carbonate crystals. *Journal of Microscopy*, 212(3):280 – 291, 2003.

[Blohm2007] D. Blohm, J. Zeng, M. Fritz, and G. Grathwohl. WO2007EP54252 20070502. 2007, WO 2007125127 (A2), 2007.

[Blum1987] H. Blum, H. Beier and H. J. Gross. Improved silver staining of plant proteins, rna and dna in polyacrylamide gels. *Electrophoresis*, 8:93 – 99, 1987.

[Bonza2000] M. C. Bonza, P. Morandini, L. Luoni, M. Geisler, M. G. Palmgren, and M. I. De Michelis. At-aca8 encodes a plasma membrane-localized calcium-atpase of arabidopsis with a calmodulin-binding domain at the n terminus. *Plant Physiology*, 123(4):1495 – 1506, 2000.

[Boudart2005] G. Boudart, E. Jamet, M. Rossignol, C. Lafitte, G. Borderies, A. Jauneau, M.-T. Esquerré-Tugayé, and R.l Pont-Lezica. Cell wall proteins in apoplastic fluids of arabidopsis thaliana rosettes: Identification by mass spectrometry and bioinformatics. *Proteomics*, 5(1):212 – 221, 2005.

[Bouropoulos2001] N. Bouropoulos, S. Weiner, and L. Addadi. Calcium oxalate crystals in tomato and tobacco plants: Morphology and in vitro interactions of crystal-associated macromolecules. *Chemistry*, 7(9):1881 – 1888, 2001.

[Brett1996] C. Brett and K. Waldron. *Biochemistry and physiology of the plant cell wall*. page 276, 1996.

[Briskin1990] D. P. Briskin. Ca^{2+}-translocating atpase of the plant plasma membrane. *Plant Physiology*, 94(2):397 – 400, 1990.

[Brown1996] R. M. J. R. Brown, I. M. Saxena, and K. Kudlicka. Cellulose biosynthesis in higher plants. *Trends in Plant Science*, 1(5):149 – 156, 1996.

[Bukau2006] B. Bukau, J. Weissman, and A. Horwich. Molecular chaperones and protein quality control. *Cell*, 125(3):443 – 451, 2006.

[Cannon2008] M. C. Cannon, K. Terneus, Q. Hall, L. Tan, Y. Wang, B. L. Wegenhart, L. Chen, D. T. A. Lamport, Y. Chen, and M. J. Kieliszewski. Self-assembly of the plant cell wall requires an extensin scaffold. *Proceedings of the National Academy of Sciences*, 105(6):2226 – 2231, 2008.

[Carpita1993] N. C. Carpita and D. M. Gibeaut. Structural models of primary cell walls in flowering plants: Consistency of molecular structure with the physical properties of the walls during growth. *The Plant Journal*, 3(1):1 – 30, 1993.

[Cavalier2008] D. M. Cavalier, O. Lerouxel, L. Neumetzler, K. Yamauchi, A. Reinecke, G. Freshour, O. A. Zabotina, M. G. Hahn, I. Burgert, M. Pauly, N. V. Raikhel, and K. Keegstra. Disrupting two arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *The Plant Cell Online*, 20(6):1519 – 1537, 2008.

[Chen2011] C.-L. Chen, J. Qi, R. N. Zuckermann, and J. J. DeYoreo. Engineered biomimetic polymers as tunable agents for controlling $caco_3$ mineralization. *Journal of the American Chemical Society*, 133(14):5214 – 5217, 2011.

[Chikwamba2003] R. K. Chikwamba, M. P. Scott, L. B. Mejía, H. S. Mason, and K. Wang. Localization of a bacterial protein in starch granules of transgenic maize kernels. *Proceedings of the National Academy of Sciences*, 100(19):11127 – 11132, 2003.

[Choi2005] Y.-E. Choi and E. Harada. Roles of calcium and cadmium on cd-containing intra- and extracellular formation of ca crystals in tobacco. *Journal of Plant Biology*, 48:113 – 119, 2005.

[Clough1998] S. J. Clough and A. F. Bent. Floral dip: A simplified method for agrobacterium-mediated transformation of arabidopsis thaliana. *The Plant Journal*, 16(6):735 – 743, 1998.

[Collino2008] S. Collino and J. S. Evans. Molecular specifications of a mineral modulation sequence derived from the aragonite-promoting protein n16. *Biomacromolecules*, 9(7):1909 – 1918, 2008.

[Conn2011] S. J. Conn, M. Gilliham, A. Athman, A. W. Schreiber, U. Baumann, I. Moller, N.-H. Cheng, M. A. Stancombe, K. D. Hirschi, A. A. Webb, R. Burton, B. N. Kaiser, S. D. Tyerman, and R. A. Leigh. Cell-specific vacuolar calcium storage mediated by cax1 regulates apoplastic calcium concentration, gas exchange, and plant productivity in arabidopsis. *The Plant Cell Online*, 23(1):240 – 257, 2011.

[Cosgrove1997] D. J. Cosgrove. Assembly and enlargement of the primary cell wall in plants. *Annual Review of Cell and Developmental Biology*, 13(1):171 – 201, 1997.

[Cosgrove2005] D. J. Cosgrove. Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology*, 6(11):850 – 861, 2005.

[Currey1974] J. D. Currey and J. D. Taylor. The mechanical behaviour of some molluscan hard tissues. *Journal of Zoology*, 173(3):395 – 406, 1974.

[Curtis2003] M. D. Curtis and U. Grossniklaus. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiology*, 133(2):462 – 469, 2003.

[Deville2006] S. Deville, E. Saiz, R. K. Nalla, and A. P. Tomsia. Freezing as a path to build complex composites. *Science*, 311(5760):515 – 518, 2006.

[Dobiasova2003] L. Dobiášová, R. Kužel, H. Šíchová, and J. Kopeček. The egg-shell microstructure studied by powder diffraction. *http://www.xray.cz/epdic/abstracts/293.html*, 2003.

[Dodenhof2010] T. Dodenhof, M. Fritz, S. Kelm, and F. Dietz. Embl accession fn674445.1, direct submission. *www.ncbi.nlm.nih.gov/protein/CBK19535.1*, 2010.

[Ellis2008] A. L. Ellis. *Rational design of calcium biosensors*. PhD thesis, Chemistry Dissertations. Paper 25. *http* : *//digitalarchive.gsu.edu/chemistry_diss*/25, 2008.

[Evans2001] A. G. Evans, Z. Suo, R. Z. Wang, I. A. Aksay, M. Y. He, and J. W. Hutchinson. Model for the robust mechanical behavior of nacre. *Journal of Materials Research*, 16:2475 – 2484, 2001.

[Falini1996] G. Falini, S. Albeck, S. Weiner, and L. Addadi. Control of aragonite or calcite polymorphism by mollusk shell macromolecules. *Science*, 271(5245):67 – 69, 1996.

[Fincher1983] G. B. Fincher, B. A. Stone, and A. E. Clarke. Arabinogalactan-proteins: Structure, biosynthesis and function. *Annual Review of Plant Physiology*, 34(1):47 – 70, 1983.

[Foster1956] A. S. Foster. Plant idioblasts: Remarkable examples of cell specialization. *Protoplasma*, 46:184 – 193, 1956.

[Fowler1999] T. J. Fowler, C. Bernhardt, and M. L. Tierney. Characterization and expression of four proline-rich cell wall protein genes in arabidopsis encoding two distinct subsets of multiple domain proteins. *Plant Physiology*, 121(4):1081 – 1091, 1999.

[Franceschi2005] V. R. Franceschi and P. A. Nakata. Calcium oxalate in plants: Formation and function. *Annual Review of Plant Biology*, 56:41 – 71, 2005.

[Franco1996] T. T. Franco, N. R. Rodrigues, G. E. Serra, V. R. Panegassi, and M. S. Buckeridge. Characterization of storage cell wall polysaccharides from brazilian legume seeds and the formation of aqueous two-phase systems. *Journal of Chromatography B: Biomedical Sciences and Applications*, 680:255–261, 1996.

[Freeman2010] C. L. Freeman, J. H. Harding, D. Quigley, and P. M. Rodger. Structural control of crystal nuclei by an eggshell protein. *Angewandte Chemie International Edition*, 49(30):5135 – 5137, 2010.

[Freeman2012] C. L. Freeman, J. H. Harding, D. Quigley, and P. M. Rodger. Protein binding on stepped calcite surfaces: Simulations of ovocleidin-17 on calcite 31.16 and 31.8. *Physical Chemistry Chemical Physics*, 14:7287 – 7295, 2012.

[Frenzel2011] M. Frenzel and E. M. Harper. Micro-structure and chemical composition of vateritic deformities occurring in the bivalve corbicula fluminea (müller, 1774). *Journal of Structural Biology*, 174(2):321 – 332, 2011.

[Fuhrmann1999] M. Fuhrmann, W. Oertel, and P. Hegemann. A synthetic gene coding for the green fluorescent protein (gfp) is a versatile reporter in chlamydomonas reinhardtii. *The Plant Journal*, 19(3):353 – 361, 1999.

[Gabius2004] H.-J. Gabius, H.-C. Siebert, S. André, J. Jiménez-Barbero, and H. Rüdiger. Chemical biology of the sugar code. *ChemBioChem*, 5(6):740 – 764, 2004.

[Gebauer2008] D. Gebauer, A. Völkel, and H. Cölfen. Stable prenucleation calcium carbonate clusters. *Science*, 322(5909):1819 – 1822, 2008.

[Gelvin2005] S. B. Gelvin. Agricultural biotechnology: Gene exchange by design. *Nature*, 433(7026):583 – 584, 2005.

[George2008] L. George, S. M. Romanowsky, J. F. Harper, and R. A. Sharrock. The aca10 ca^{2+}-atpase regulates adult vegetative development and inflorescence architecture in arabidopsis. *Plant Physiology*, 146(2):716–728, 2008.

[Gibeaut1991] D. M. Gibeaut and N. C. Carpita. Tracing cell wall biogenesis in intact cells and plants. *Plant Physiology*, 97(2):551 – 561, 1991.

[Giddings1980] T. H. Giddings, D. L. Brower, and L. A. Staehelin. Visualization of particle complexes in the plasma membrane of micrasterias denticulata associated with the formation of cellulose fibrils in primary and secondary cell walls. *The Journal of Cell Biology*, 84(2):327 – 339, 1980.

[Goldberg1996] R. Goldberg, C. Morvan, A. Jauneau, and M. C. Jarvis. *Pectins and Pectinases. Methyl-esterification, de-esterification and gelation of pectins in the primary cell wall.* Elsevier Science, Amsterdam, pages 151 - 171, 1996.

[Gotliv2003] B.-A. Gotliv, L. Addadi, and S. Weiner. Mollusk shell acidic proteins: In search of individual functions. *ChemBioChem*, 4(6):522 – 529, 2003.

[Harada2008] E. Harada and Y. E. Choi. Investigation of metal exudates from tobacco glandular trichomes under heavy metal stresses using a variable pressure scanning electron microscopy system. *Plant Biotechnology*, 25(4):407 – 411, 2008.

[Harrison2006] S. Harrison, E. Mott, K. Parsley, S. Aspinall, J. Gray, and A. Cottage. A rapid and robust method of identifying transformed arabidopsis thaliana seedlings following floral dip transformation. *Plant Methods*, 2(1):19, 2006.

[Hayashi1989] T. Hayashi. Xyloglucans in the primary cell wall. *Annual Review of Plant Physiology and Plant Molecular Biology*, 40(1):139 – 168, 1989.

[Heinemann2006] F. Heinemann, L. Treccani, and M. Fritz. Abalone nacre insoluble matrix induces growth of flat and oriented aragonite crystals. *Biochemical and Biophysical Research Communications*, 344(1):45 – 49, 2006.

[Heinemann2011] F. Heinemann, M. Gummich, M. Radmacher, and M. Fritz. Modification of $caco_3$ precipitation rates by water-soluble nacre proteins. *Materials Science and Engineering: C*, 31(2):99 – 105, 2011.

[Hincke1995] M. Hincke, C. Tsang, M. Courtney, V. Hill, and R. Narbaitz. Purification and immunochemistry of a soluble matrix protein of the chicken eggshell (ovocleidin 17). *Calcified Tissue International*, 56:578 – 583, 1995.

[Hood1999] E. E. Hood and J. M. Jilka. Plant-based production of xenogenic proteins. *Current Opinion in Biotechnology*, 10(4):382 – 386, 1999.

[Hood2002] E. E. Hood. From green plants to industrial enzymes. *Enzyme and Microbial Technology*, 30(3):279 – 283, 2002.

[Hood2004] E. E. Hood. Bioindustrial and biopharmaceutical products from plants. In Fischer T. et al., editor, *New directions for a diverse planet: Proceedings for the 4th International Crop Science Congress*, Brisbane, Australia, 2004.

[Hudgins2003] J. W. Hudgins, T. Krekling, and V. R. Franceschi. Distribution of calcium oxalate crystals in the secondary phloem of conifers: A constitutive defense mechanism? *New Phytologist*, 159(3):677 – 690, 2003.

[Huettner2012] S. Hüttner and R. Strasser. Endoplasmic reticulum-associated degradation of glycoproteins in plants. *Frontiers in Plant Science - Plant Cell Biology*, 67(3):1–6, 2012.

[Isaure2010] M.-P. Isaure, G. Sarret, E. Harada, Y.-E. Choi, M. A. Marcus, S. C. Fakra, N. Geoffroy, S. Pairis, J. Susini, S. Clemens, and A. Manceau. Calcium promotes cadmium elimination as vaterite grains by tobacco trichomes. *Geochimica et Cosmochimica Acta*, 74(20):5817 – 5834, 2010.

[Kahr2001] B. Kahr and R. W. Gurney. Dyeing crystals. *Chemical Reviews*, 101(4):893 – 952, 2001.

[Katayama2008] H. Katayama, Y. Fujibayashi, S. Nagaoka, and Y. Sugimura. Ultrastructural and immunochemical features of the cell wall sac formed in mulberry (morus alba) idioblasts. *Journal of Plant Research*, 121:201 – 205, 2008.

[Keene2010] E. C. Keene, J. S. Evans, and L. A. Estroff. Silk fibroin hydrogels coupled with the n16n beta-chitin complex: An in vitro organic matrix for controlling calcium carbonate mineralization. *Crystal Growth and Design*, 10(12):5169 – 5175, 2010.

[Kiermayer1979] O. Kiermayer and U. B. Sleytr. Hexagonally ordered rosettes of particles in the plasma membrane of micrasterias denticulata and their significance for microfibril formation and orientation. *Protoplasma*, 101:133 – 138, 1979.

[Kim2006] S. A. Kim, T. Punshon, A. Lanzirotti, L. Li, J. M. Alonso, J. R. Ecker, J. Kaplan, and M. L. Guerinot. Localization of iron in arabidopsis seed requires the vacuolar membrane transporter vit1. *Science*, 314(5803):1295 – 1298, 2006.

[Klatte2009] M. Klatte and P. Bauer. Accurate real-time reverse transcription quantitative pcr. *Plant Signal Transduction*, 479:1 – 17, 2009.

[Kristensen2008] J. Kristensen, L. Thygesen, C. Felby, H. Jorgensen, and T. Elder. Cell-wall structural changes in wheat straw pretreated for bioethanol production. *Biotechnology for Biofuels*, 1(1):5, 2008.

[Kwon2008] C. Kwon, C. Neu, S. Pajonk, H. S. Yun, U. Lipka, M. Humphry, S. Bau, M. Straus, M. Kwaaitaal, H. Rampelt, F. E. Kasmi, G. Jurgens, J. Parker, R. Panstruga, V. Lipka, and P. Schulze-Lefert. Co-option of a default secretory pathway for plant immune responses. *Nature*, 451(7180):835 – 840, 2008.

[Lakshminarayanan2005] R. Lakshminarayanan, E. O. Chi-Jin, X. J. Loh, R. M. Kini, and S. Valiyaveettil. Purification and characterization of a vaterite-inducing peptide, pelovaterin, from the eggshells of pelodiscus sinensis (chinese soft-shelled turtle). *Biomacromolecules*, 6(3):1429 – 1437, 2005.

[Laraia1989] V. J. Laraia and A. H. Heuer. Novel composite microstructure and mechanical behavior of mollusk shell. *Journal of the American Ceramic Society*, 72(11):2177 – 2179, 1989.

[Lee1977] S. L. Lee, A. Veis, and T. Glonek. Dentin phosphoprotein: An extracellular calcium-binding protein. *Biochemistry*, 16(13):2971 – 2979, 1977.

[Levi-Kalisman2001] Y. Levi-Kalisman, G. Falini, L. Addadi, and S. Weiner. Structure of the nacreous organic matrix of a bivalve mollusk shell examined in the hydrated state using cryo-tem. *Journal of Structural Biology*, 135(1):8 – 17, 2001.

[Levi1998] Y. Levi, S. Albeck, A. Brack, S. Weiner, and L. Addadi. Control over aragonite crystal nucleation and growth: An in vitro study of biomineralization. *Chemistry - A European Journal*, 4(3):389 – 396, 1998.

[Lionetti2010] V. Lionetti, F. Francocci, S. Ferrari, C. Volpi, D. Bellincampi, R. Galletti, R. D'Ovidio, G. De Lorenzo, and F. Cervone. Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. *Proceedings of the National Academy of Sciences*, 107(2):616 – 621, 2010.

[Lippmann1973] F. Lippmann. *Sedimentary carbonate minerals*. Springer Berlin / Heidelberg, pages 116 - 131, 1973.

[Loris2002] R. Loris. Principles of structures of animal and plant lectins. *Biochim Biophys Acta*, 1572(2-3):198 – 208, 2002.

[Lowenstam1981] H. A. Lowenstam. Minerals formed by organisms. *Science*, 211(4487):1126 – 1131, 1981.

[Lowenstam1989] H. A. Lowenstam and S. Weiner. *On biomineralization*. Oxford University Press, New York, 1989.

[Ma2003] J. K-C. Ma, P. M. W. Drake, and P. Christou. The production of recombinant pharmaceutical proteins in plants. *Nature Reviews Genetics*, 4(10):794 – 805, 2003.

[Macquet2007] A. Macquet, M.-C. Ralet, J. Kronenberger, A. Marion-Poll, and H. M. North. In situ, chemical and macromolecular study of the composition of arabidopsis thaliana seed coat mucilage. *Plant and Cell Physiology*, 48(7):984 – 999, 2007.

[Malhotra2007] J. D. Malhotra and R. J. Kaufman. The endoplasmic reticulum and the unfolded protein response. *Seminars in Cell and Developmental Biology*, 18(6):716 – 731, 2007.

[Mann1983] P. Connett, H. Follmann, M. Lammers, S. Mann, J. Odom, K. Wetterhahn, and Stephen Mann. Mineralization in biological systems. In *Inorganic elements in biochemistry*, volume 54 of *Structure and Bonding*, pages 125 – 174. Springer Berlin, Heidelberg, 1983.

[Mann2000] K. Mann, I. M. Weiss, S. André, H. J. Gabius, and M. Fritz. The amino-acid sequence of the abalone (haliotis laevigata) nacre protein perlucin. *European Journal of Biochemistry*, 267(16):5257 – 5264, 2000.

[Mann2001] S. Mann. *Biomineralization. Principles and concepts in bioinorganic materials chemistry*. Oxford Univiversity Press, New York, 2001.

[Marin-Garcia2008] L. Marín-García, B. A. Frontana-Uribe, J. P. , Reyes-Grajeda, V. Stojanoff, H. J. Serrano-Posada, and A. Moreno. Chemical recognition

of carbonate anions by proteins involved in biomineralization processes and their influence on calcite crystal growth. *Crystal Growth and Design*, 8(4):1340 – 1345, 2008.

[Mayer2005] G. Mayer. Rigid biological systems as models for synthetic composites. *Science*, 310(5751):1144 – 1147, 2005.

[Metzler2010] R. A. Metzler, J. S. Evans, C. E. Killian, D. Zhou, T. H. Churchill, N. P. Appathurai, S. N. Coppersmith, and P. U. P. A. Gilbert. Nacre protein fragment templates lamellar aragonite growth. *Journal of the American Chemical Society*, 132(18):6329 – 6334, 2010.

[Meyer2009] D. Meyer, S. Pajonk, C. Micali, R. O'Connell, and P. Schulze-Lefert. Extracellular transport and integration of plant secretory proteins into pathogen-induced cell wall compartments. *The Plant Journal*, 57(6):986 – 999, 2009.

[Micheli2001] F. Micheli. Pectin methylesterases: Cell wall enzymes with important roles in plant physiology. *Trends in Plant Science*, 6(9):414 – 419, 2001.

[Miyamoto1996] H. Miyamoto, T. Miyashita, M. Okushima, S. Nakano, T. Morita, and A. Matsushiro. A carbonic anhydrase from the nacreous layer in oyster pearls. *Proceedings of the National Academy of Sciences*, 93(18):9657 – 9660, 1996.

[Montagnani2011] C. Montagnani, B. Marie, F. Marin, C. Belliard, F. Riquet, A. Tayalé, I. Zanella-Cléon, E. Fleury, Y. Gueguen, D. Piquemal, and N. Cochennec-Laureau. Pmarg-pearlin is a matrix protein involved in nacre framework formation in the pearl oyster pinctada margaritifera. *ChemBioChem*, 12(13):2033 – 2043, 2011.

[Morikawa1978] H. Morikawa, R. Hayashi, and M. Senda. Infrared analysis of pea stem cell walls and oriented structure of matrix polysaccharides in them. *Plant and Cell Physiology*, 19(7):1151 – 1159, 1978.

[Nakata2000] P. A. Nakata and M. M. McConn. Isolation of medicago truncatula mutants defective in calcium oxalate crystal formation. *Plant Physiology*, 124(3):1097 – 1104, 2000.

[Nakata2012] P. A. Nakata. Engineering calcium oxalate crystal formation in arabidopsis. *Plant and Cell Physiology*, 2012.

[Ng1998] K. K.-S. Ng, S. Park-Snyder, and W. I. Weis. Ca^{2+}-dependent structural changes in c-type mannose-binding proteins. *Biochemistry*, 37(51):17965 – 17976, 1998.

[Nicol1998] F. Nicol, I. His, A. Jauneau, S. Vernhettes, H. Canut, and H. Höfte. A plasma membrane-bound putative endo-1,4-[beta]-d-glucanase is required for normal wall assembly and cell elongation in arabidopsis. *EMBO J*, 17(19):5563 – 5576, 1998.

[Nitta2006] I. Nitta, A. Kida, Y. Fujibayashi, H. Katayama, and Y. Sugimura. Calcium carbonate deposition in a cell wall sac formed in mulberry idioblasts. *Protoplasma*, 228:201 – 208, 2006.

[Nys2004] Y. Nys, J. Gautron, J. M. Garcia-Ruiz, and M. T. Hincke. Avian eggshell mineralization: Biochemical and functional characterization of matrix proteins. *Comptes Rendus Palevol*, 3(6 - 7):549 – 562, 2004.

[Pai2011] R. K. Pai and M. Cotlet. Highly stable, water-soluble, intrinsic fluorescent hybrid scaffolds for imaging and biosensing. *The Journal of Physical Chemistry C*, 115(5):1674 – 1681, 2011.

[Park2003] Y. W. Park, R. Tominaga, J. Sugiyama, Y. Furuta, E. Tanimoto, M. Samejima, F. Sakai, and T. Hayashi. Enhancement of growth by expression of poplar cellulase in arabidopsis thaliana. *The Plant Journal*, 33(6):1099 – 1106, 2003.

[Pauly2010] M. Pauly and K. Keegstra. Plant cell wall polymers as precursors for biofuels. *Current Opinion in Plant Biology*, 13(3):304 – 311, 2010.

[Peng2002] L. Peng, Y. Kawagoe, P. Hogan, and D. Delmer. Sitosterol-beta-glucoside as primer for cellulose synthesis in plants. *Science*, 295(5552):147 – 150, 2002.

[Ponce2011] C. B. Ponce and J. S. Evans. Polymorph crystal selection by n16, an intrinsically disordered nacre framework protein. *Crystal Growth and Design*, 11(10):4690 – 4696, 2011.

[Popescu2012] S. C. Popescu. A model for the biosynthesis and transport of plasma membrane-associated signaling receptors to the cell surface. *Frontiers in Plant Science - Plant Proteomics*, 71(3):doi: 10.3389/fpls.2012.00071, 2012.

[Punshon2012] T. Punshon, K. Hirschi, J. Yang, A. Lanzirotti, B. Lai, and M. L. Guerinot. The role of cax1 and cax3 in elemental distribution and abundance in arabidopsis seed. *Plant Physiology*, 158(1):352 – 362, 2012.

[Reyes-Grajeda2004] J. P. Reyes-Grajeda, A. Moreno, and A. Romero. Crystal structure of ovocleidin-17, a major protein of the calcified gallus gallus eggshell. *Journal of Biological Chemistry*, 279(39):40876 – 40881, 2004.

[Rizk2000] S. E. Rizk, R. M. Abdel-Massih, E. A.-H. Baydoun, and C. T. Brett. Protein- and ph-dependent binding of nascent pectin and glucuronoarabinoxylan to xyloglucan in pea. *Planta*, 211:423 – 429, 2000.

[Roemisch2005] K. Römisch. Endoplasmic reticulum-associated degradation. *Annual Review of Cell and Developmental Biology*, 21:435 – 456, 2005.

[Rose2009] M. Rose and M. Hincke. Protein constituents of the eggshell: Eggshell-specific matrix proteins. *Cellular and Molecular Life Sciences*, 66:2707 – 2719, 2009.

[Sagi2005] G. Sagi, A. Katz, D. Guenoune-Gelbart, and B. L. Epel. Class 1 reversibly glycosylated polypeptides are plasmodesmal-associated proteins delivered to plasmodesmata via the golgi apparatus. *The Plant Cell Online*, 17(6):1788 – 1800, 2005.

[Samata1999] T. Samata, N. Hayashi, M. Kono, K. Hasegawa, C. Horita, and S. Akera. A new matrix protein family related to the nacreous layer formation of pinctada fucata. *FEBS Letters*, 462(1-2):225 – 229, 1999.

[Sambrook2001] J. Sambrook and D. W. Russell. *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, 2001.

[Sanders1999] D. Sanders, C. Brownlee, and J. F. Harper. Communicating with calcium. *The Plant Cell Online*, 11(4):691 – 706, 1999.

[Sarret2006] G. Sarret, E. Harada, Y.-E. Choi, M.-P. Isaure, N. Geoffroy, S. Fakra, M. A. Marcus, M. Birschwilks, S. Clemens, and A. Manceau. Trichomes of tobacco excrete zinc as zinc-substituted calcium carbonate and other zinc-containing compounds. *Plant Physiolgy*, 141(3):1021 – 1034, 2006.

[Sato2011] A. Sato, S. Nagasaka, K. Furihata, S. Nagata, I. Arai, K. Saruwatari, T. Kogure, S. Sakuda, and H. Nagasawa. Glycolytic intermediates induce amorphous calcium carbonate formation in crustaceans. *Nature Chemical Biology*, 7(4):197 – 199, 2011.

[Scheller2001] J. Scheller, K. H. Gührs, F. Grosse, and U. Conrad. Production of spider silk proteins in tobacco and potato. *Nature Biotechnology*, 19:573 – 577, 2001.

[Showalter1993] A. M. Showalter. Structure and function of plant cell wall proteins. *Plant Cell*, 5(1):9 – 23, 1993.

[Spann2010] N. Spann, E. Harper, and D. Aldridge. The unusual mineral vaterite in shells of the freshwater bivalve corbicula fluminea from the uk. *Naturwissenschaften*, 97:743 – 751, 2010.

[Stein2009] H. Stein, M. Wilensky, Y. Tsafrir, M. Rosenthal, R. Amir, T. Avraham, K. Ofir, O. Dgany, A. Yayon, and O. Shoseyov. Production of bioactive, post-translationally modified, heterotrimeric, human recombinant type-i collagen in transgenic tobacco. *Biomacromolecules*, 10(9):2640 – 2645, 2009.

[Streatfield2007] S. J. Streatfield. Approaches to achieve high-level heterologous protein production in plants. *Plant Biotechnology Journal*, 5(1):2– 15, 2007.

[Sugimura1998] Y. Sugimura, I. Nitta, Y. Morita, S. Ishikawa, T. Mori, E. Kotani, and T. Furusawa. Microscopic detection of calcium deposited in idioblasts of mulberry leaves. *Journal of Sericultural Science of Japan*, 67:445 – 451, 1998.

[Sugimura1999] Y. Sugimura, T. Mori, I. Nitta, E. Kotani, T. Furusawa, M. Tatsumi, S. I. Kusakari, M. Wada, and Y. Morita. Calcium deposition in idioblasts of mulberry leaves. *Annals of Botany*, 83(5):543 – 550, 1999.

[Sugimura2007] Y. Sugimura and I. Nitta. Cytological changes during cell wall sac formation in mulberry idioblasts. *Protoplasma*, 231:123 – 125, 2007.

[Suzuki2009] M. Suzuki, K. Saruwatari, T. Kogure, Y. Yamamoto, T. Nishimura, T. Kato, and H. Nagasawa. An acidic matrix protein, pif, is a key macromolecule for nacre formation. *Science*, 325(5946):1388 – 1390, 2009.

[Tang2003] Z. Tang, N. A. Kotov, S. Magonov, and B. Ozturk. Nanostructured artificial nacre. *Nature Materials*, 2(6):413 – 418, 2003.

[Tian2004] G.-W. Tian, A. Mohanty, S. N. Chary, S. Li, B. Paap, G. Drakakaki, C. D. Kopec, J. Li, D. Ehrhardt, D. Jackson, S. Y. Rhee, N. V. Raikhel, and V. Citovsky. High-throughput fluorescent tagging of full-length arabidopsis gene products in planta. *Plant Physiology*, 135(1):25 – 38, 2004.

[Urade2007] R. Urade. Cellular response to unfolded proteins in the endoplasmic reticulum of plants. *FEBS Journal*, 274(5):1152 – 1171, 2007.

[Vasilevski2012] A. Vasilevski, F. A. Giorgi, L. Bertinetti, and B. Usadel. Lasso modeling of the arabidopsis thaliana seed/seedling transcriptome: A model case for detection of novel mucilage and pectin metabolism genes. *Molecular BioSystems*, DOI: 10.1039/c2mb25096a, 2012.

[Veinott2009] G. I. Veinott, T. R. Porter, and L. Nasdala. Using mg as a proxy for crystal structure and sr as an indicator of marine growth in vaterite and aragonite otoliths of aquaculture rainbow trout. *Transactions of the American Fisheries Society*, 138(5):1157 – 1165, 2009.

[Vincken2003] J. P. Vincken, H. A. Schols, R. J. F. J. Oomen, M. C. McCann, P. Ulvskov, A. G. J. Voragen, and R. G. F. Visser. If homogalacturonan were a side chain of rhamnogalacturonan i. implications for cell wall architecture. *Plant Physiology*, 132(4):1781 – 1789, 2003.

[Vitale2008] A. Vitale and R. S. Boston. Endoplasmic reticulum quality control and the unfolded protein response: Insights from plants. *Traffic*, 9(10):1581 – 1588, 2008.

[Vlasak2009] J. Vlasak, M. C. Bussat, S. Wang, E. Wagner-Rousset, M. Schaefer, C. Klinguer-Hamour, M. Kirchmeier, N. Corvaïa, R. Ionescu, and A. Beck. Identification and characterization of asparagine deamidation in the light chain cdr1 of a humanized igg1 antibody. *Analytical Biochemistry*, 392(2):145 – 154, 2009.

[Wang2008] N. Wang, Y. H. Lee, and J. Lee. Recombinant perlucin nucleates the growth of calcium carbonate crystals: Molecular cloning and characterization of perlucin from disk abalone, haliotis discus discus. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 149(2):354 – 361, 2008.

[Weber2012] E. Weber, C. Guth, M. Eder, P. Bauer, E. Arzt, and I. M. Weiss. Biotechnological mineral composites via vaterite precursors. *MRS Online Proceedings Library*, 1465:mrss12-1465-ss1403-1410 doi:1410.1557/opl.2012.1046., 2012.

[Weiner2009] S. Weiner, J. Mahamid, Y. Politi, Y. Ma, and L. Addadi. Overview of the amorphous precursor phase strategy in biomineralization. *Frontiers of Materials Science in China*, 3:104 – 108, 2009.

[Weiss2000] I. M. Weiss, S. Kaufmann, K. Mann, and M. Fritz. Purification and characterization of perlucin and perlustrin, two new proteins from the shell of the mollusc haliotis laevigata. *Biochemical and Biophysical Research Communications*, 267(1):17 – 21, 2000.

[Weiss2006] I. M. Weiss and V. Schönitzer. The distribution of chitin in larval shells of the bivalve mollusk mytilus galloprovincialis. *Journal of Structural Biology*, 153(3):264 – 277, 2006.

[Weiss2008] I. M. Weiss and F. Marin. The role of enzymes in biomineralization processes. In *Biomineralization*, pages 71 – 126. John Wiley and Sons, Ltd, 2008.

[Wheeler1981] A. P. Wheeler, J. W. George, and C. A. Evans. Control of calcium carbonate nucleation and crystal growth by soluble matrix of oyster shell. *Science*, 212(4501):1397 – 1398, 1981.

[Willats2001] W. G. T. Willats, L. McCartney, W. Mackie, and J. P. Knox. Pectin: Cell biology and prospects for functional analysis. *Plant Molecular Biology*, 47:9 – 27, 2001.

[Wong1984] M. Wong, M. J. C. Hendrix, K. von der Mark, C. Little, and R. Stern. Collagen in the egg shell membranes of the hen. *Developmental Biology*, 104(1):28 – 36, 1984.

[Worms1986] D. Worms and S. Weiner. Mollusk shell organic matrix: Fourier transform infrared study of the acidic macromolecules. *Journal of Experimental Zoology*, 237(1):11 – 20, 1986.

[Worrall1992] D. Worrall, D. L. Hird, R. Hodge, W. Paul, J. Draper, and R. Scott. Premature dissolution of the microsporocyte callose wall causes male sterility in transgenic tobacco. *Plant Cell*, 4(7):759 – 771, 1992.

[Wyman2005] C. E. Wyman, B. E. Dale, R. T. Elander, M. Holtzapple, M. R. Ladisch, and Y. Y. Lee. Coordinated development of leading biomass pretreatment technologies. *Bioresource Technology*, 96(18):1959 – 1966, 2005.

[Xu2011] X. Xu, Q. Gan, R. Clough, K. Pappu, J. Howard, J. Baez, and K. Wang. Hydroxylation of recombinant human collagen type i alpha 1 in transgenic maize co-expressed with a recombinant human prolyl 4-hydroxylase. *BMC Biotechnology*, 11(1):69, 2011.

[Yang1996] F. Yang, L. G. Moss, and G. N. Phillips. The molecular structure of green fluorescent protein. *Nature Biotechnology*, 14(10):1246 – 1251, 1996.

[Yang2005] J. Yang, L. A. Barr, S. R. Fahnestock, and Z.-B. Liu. High yield recombinant silk-like protein production in transgenic plants through protein targeting. *Transgenic Research*, 14:313 – 324, 2005.

[Ziegler2000] M. T. Ziegler, S. R. Thomas, and K. J. Danna. Accumulation of a thermostable endo-1,4-beta-d-glucanase in the apoplast of arabidopsis thaliana leaves. *Molecular Breeding*, 6(1):37 – 46, 2000.

[Zykwinska2005] A. W. Zykwinska, M. C. J. Ralet, C. D. Garnier, and J. F. J. Thibault. Evidence for in vitro binding of pectin side chains to cellulose. *Plant Physiology*, 139(1):397 – 407, 2005.

5 Appendix

5.1 DNA Sequences

Perlucin

```
  1 GGATGTCCTT TGGGTTTTCA CCAACATCGT CGAAGCTGCT ATTGGTTCTC
 51 AACCATCAAG TCCTCATTTG CTGAAGCTGC TGGTTACTGT AGATACCTAG
101 AGTCGCATTT GGCGATCATA TCCAACAAAG ACGAAGATAG CTTCATTAGA
151 GGCTATGCAA CAAGGCTTGG CGAAGCATTC AACTATTGGC TTGGTGCAAG
201 CGATCTGAAT ATAGAGGGAA GATGGTTATG GGAGGGACAA CGAAGGATGA
251 ACTATACCAA TTGGAGTCCG GGACAACCAG ATAACGCTGG GGGAATTGAA
301 CATTGCTTGG AGCTTAGACG TGATCTTGGG AATTACCTCT GGAATGACTA
351 CCAGTGTCAA AAGCCATCAC ATTTCATCTG CGAGAAGGAA AGGATTCCTT
401 ATACGAATTC CCTCCACGCA AATCTACAGC AAAGAGATTC GCTTCACGCC
451 AATCTGCAAC AAAGA
```

N16N

```
  1 GCATACCACA AGAAATGCGG AAGATATAGC TATTGCTGGA TACCGTACGA
 51 CATCGAAAGG GATAGATACG ACAACGGGGA TAAAAAGTGC
```

Ovocleidin-17 (OC17)

```
  1 GACCCTGACG GCTGTGGACC TGGATGGGTC CCAACCCCGG GTGGCTGCCT
 51 CGGCTTCTTT AGCCGGGAAC TTAGCTGGAG CCGTGCCGAG AGCTTTTGCC
101 GACGTTGGGG ACCTGGATCA CATCTGGCCG CTGTCCGGAG CGCGGCTGAA
151 TTGAGACTTC TCGCCGAACT GCTCAACGCC TCAAGAGGAG GCGATGGATC
201 TGGGGAAGGG GCTGATGGGA GAGTGTGGAT TGGCCTGCAT CGACCAGCAG
251 GAAGCCGATC TTGGAGATGG AGTGACGGGA CAGCTCCGCG TTTTGCTTCC
301 TGGCACCGAA CTGCAAAGGC ACGAAGAGGA GGCAGGTGCG CTGCACTAAG
351 GGACGAAGAG GCCTTTACAA GCTGGGCGGC AAGACCGTGC ACGGAAAGGA
401 ACGCCTTCGT TTGCAAAGCG GCGGCG
```

GFP pGFP-Sca-CDB-8, pQE System as vector backbone

```
  1 ATGAGAGGAT CTCACCATCA CCATCACCAT ACGGATCCCG CCAAGGGCGA
 51 GGAGCTGTTC ACCGGTGTGG TCCCCATCCT GGTGGAGCTG GACGGCGACG
101 TGAACGGCCA CAAGTTCTCC GTCTCCGGCG AGGGTGAGGG TGACGCCACC
151 TACGGCAAGC TGACCCTGAA GTTCATCTGC ACCACCGGCA AGCTGCCCGT
201 GCCCTGGCCC ACCCTGGTCA CCACCCTGAC CTACGGTGTG CAGTGCTTCT
251 CCCGCTACCC CGACCACATG AAGCAGCACG ACTTCTTCAA GTCCGCCATG
301 CCCGAGGGCT ACGTGCAGGA GCGCACCATC TTCTTCAAGG ACGACGGCAA
351 CTACAAGACC CGCGCCGAGG TCAAGTTCGA GGGCGACACC CTGGTGAACC
401 GCATCGAGCT GAAGGGCATC GACTTCAAGG AGGACGGCAA CATCCTGGGC
451 CACAAGCTGG AGTACAACTA CAACTCCCAC AACGTGTACA TCATGGCCGA
501 CAAGCAGAAG AACGGCATCA AGGTGAACTT CAAGATCCGC CACAACATCG
551 AGGACGGCTC CGTGCAGCTG GCCGACCACT ACCAGCAGAA CACCCCCATC
601 GGCGATGGCC CCGTGCTGCT GCCCGACAAC CACTACCTGT CCACCCAGTC
651 CGCCCTGTCC AAGGACCCCA ACGAGAAGCG CGACCACATG GTCCTGCTGG
701 AGTTCGTCAC CGCTGCCGGC ATCACCCACG GCATGGACGA GCTGTACAAG
751 GGAGGCGGGG CAGCCAAGCT TAATTA
```

Signal peptide EXT3 (SP-$_{EXT3}$, AT1G21310)

```
  1 ATGGGTTCTC CTATGGCCTC TTTAGTTGCT ACTCTCCTAG TGCTTACGAT
 51 TTCGCTGACA TTCGTCTCTC AATCTACTGC T
```

Signal peptide PRP4 (SP-$_{\text{PRP4}}$, AT4G38770)

```
  1 ATGCGCATCT TACCTGAACC TAGAGGTAGT GTGCCCTGTC TCCTATTGCT
 51 CGTATCTGTT CTTCTGTCTG CTACTCTTTC TCTTGCT
```

Signal peptide CH (SP-$_{\text{Chitin-hydrolase}}$, AT4G19810)

```
  1 ATGTCATCTA CAAAACTGAT TAGCCTGATT GTCTCTATTA CTTTTTTTCT
 51 CACTCTCCAA TGTTCTATGG CTCAAACTGT AGTTAAGGCT
```

All Signal peptides were fused to "CACC" sequence for a facilitated subcloning procedure to pENTR/D-TOPO vector system.

5.2 Expression of Biomineralization Proteins in Bacteria

5.2.1 Purification Steps of Recombinant Proteins

A comparison of purification steps under standard and optimized conditions of recombinant produced proteins is shown in Figure 5.1. Cell pellets of GFP, GFP-perlucin and perlucin expressing cells were captured before cell lysis (A) and after cell lysis (B). Here, the cell pellet of GFP expressing cells were fluorescent before lysis (A) and not fluorescent after the cell lysis (B), whereas the cell pellet of GFP-perlucin expressing cells remained still fluorescent after cell lysis (compare A and B). The supernatant after cell lysis (C) and the eluate (D) was fluorescent for GFP and for GFP-perlucin only after optimized conditions.

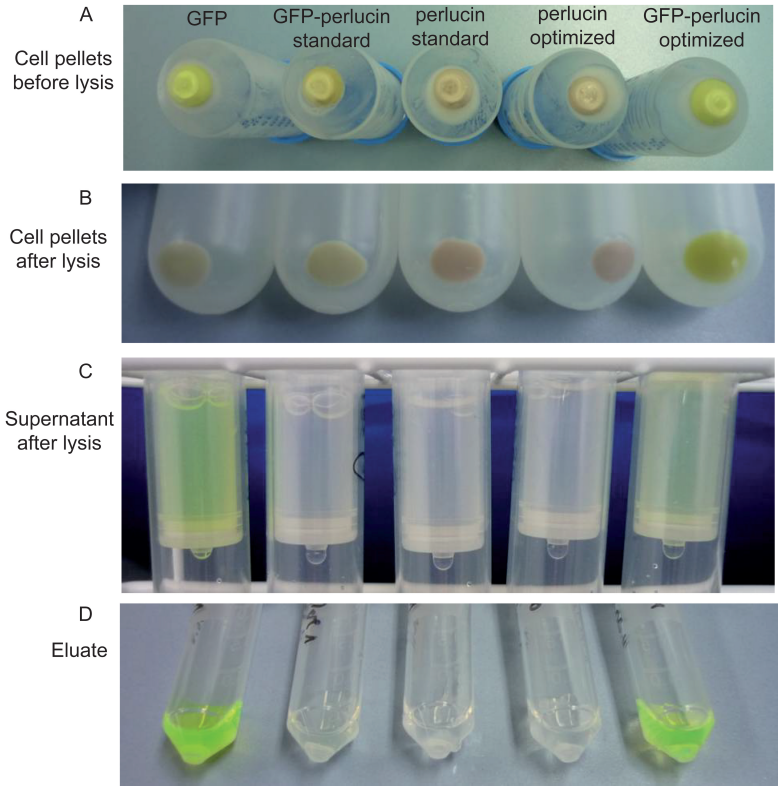


Figure 5.1: Protein purification steps of GFP, perlucin and GFP-perlucin. Cell pellets before lysis (A), after lysis (B), supernatant after lysis (C) and eluate steps (D) are shown under standard and optimized conditions. The GFP-perlucin expression is stronger under optimized conditions compared to standard conditions (D, fluorescent elutate). A part of the protein remained in the GFP-perlucin pellet fraction after lysis (B, fluorescent pellet) compared to GFP (B, no fluorescent pellet).

5.2.2 SDS-PAGE Analysis of Recombinant Proteins

SDS-PAGE analysis of expression and native Ni-NTA-affinity purification of GFP (28.9 kDa), GFP-perlucin (46.8 kDa) and perlucin (20.2 kDa) under standard and optimized conditions is shown in Figure 5.2 (standard and optimized conditions were defined in the result section 3.3). The analysis revealed that GFP was highly soluble under standard growth conditions compared to the GFP-perlucin, which was semi-soluble (compare also images of pellets in Figure 5.1). Here, high protein yields were only achieved under optimized conditions. Perlucin without tag was not soluble under native conditions (Figure 5.2, bottom row).

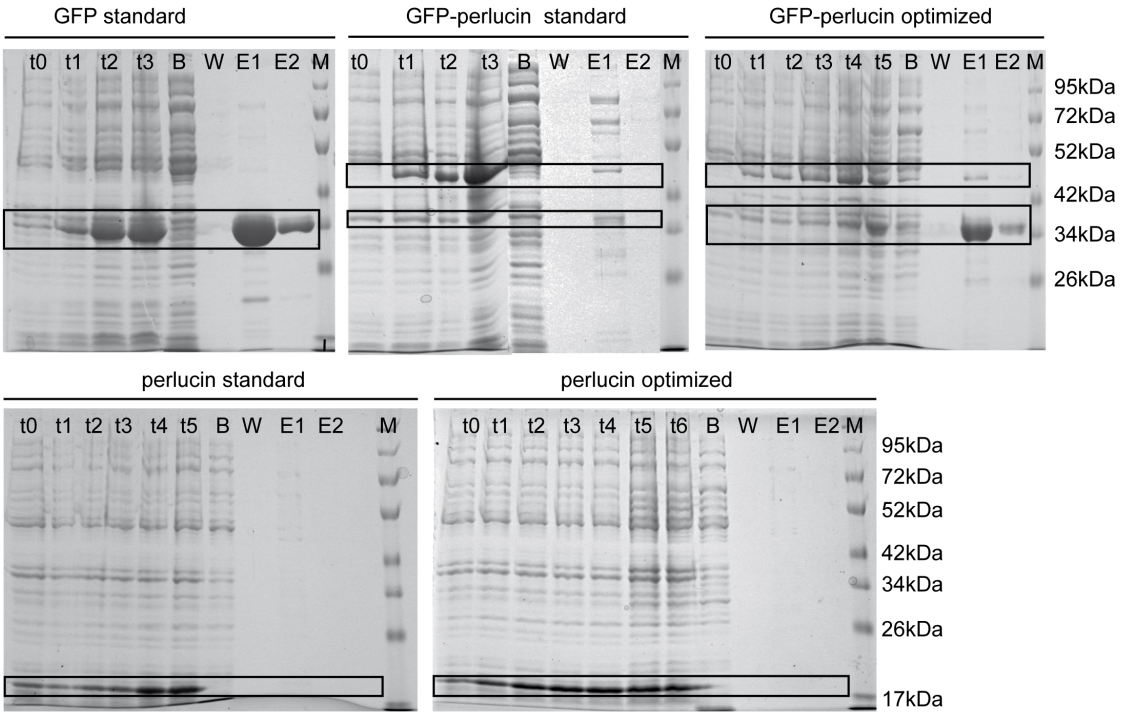


Figure 5.2: SDS-PAGE analysis (10% AA) of expression and native Ni- NTA-affinity purification of GFP (28.9 kDa), GFP-perlucin (46.8 kDa) and perlucin (20.2 kDa). Cell cultures were analyzed in 1-hour intervals (t0-t4) and on the following day (t5-t6). The binding (B), washing (W), and elution steps (E1 and E2) demonstrated the success in native affinity purification of GFP-perlucin. While GFP was easily obtained under standard growth conditions, the yield of natively purified GFP-perlucin depended largely on the culture conditions (top row). Perlucin (20.2 kDa) was expressed, but cannot be recovered in native form, irrespective of the growth conditions (bottom row). Identical molecular weight markers were used for all gels (M).

5.2.3 Western Blot Analysis of GFP-Perlucin

Western blot analysis of recombinant GFP-perlucin using anti GFP and HIS_6 antibodies revealed that the protein migrated in 3 bands on the SDS-PAGE (Figure 5.3). Both antibodies recognized identical protein fractions. According to molecular weight marker, the fragments had sizes of ∼72 kDa, ∼46 kDa and ∼42 kDa. Mass spectrometric analysis identified the 46 kDa fragment as full length variant (46.8 kDa) and the 42 kDa fragment as truncated protein variant with a molecular weight of 29.8 and 30.3 kDa (section 3.1.3.1). The upper band was not analyzed so far, but their size indicates a dimer formation of the truncated and full length variant.

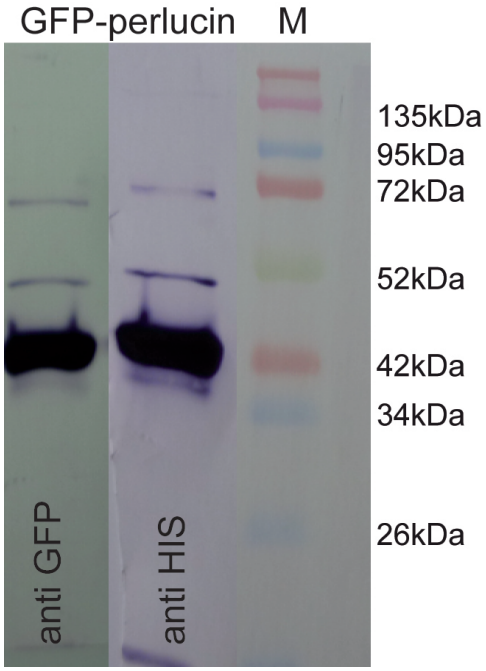


Figure 5.3: Western blot analysis of recombinant GFP-perlucin using anti GFP and HIS_6 antibodies bore a specific migration pattern of recombinant GFP-perlucin on the SDS-PAGE. Mass spectrometric analysis identified the 46 kDa fragment as full length variant (46.8 kDa) and the 42 kDa fragment as truncated protein variant of a molecular weight of 29.8 and 30.3 kDa (section 3.1.3.1). The upper band was not analyzed so far.

5.2.4 Microscopical Analysis of Calcium Carbonate Precipitates

Calcium carbonate morphologies of large volume control assay without protein additive bore rhombohedral crystals with rounded edges (Figure 5.4 A). No fluorescence signals were derived from crystal without the addition of GFP or GFP-perlucin (Figure 5.4 B). The LC-PolScope analysis in birefringent retardance (C) and orientation of the slow axis vector (D) imaging modes revealed a polycrystalline character of the minerals.

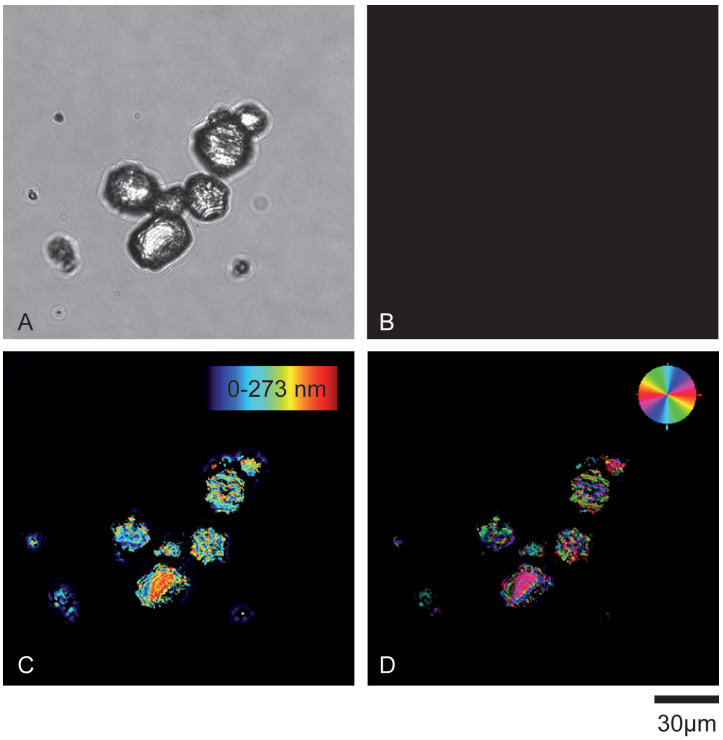


Figure 5.4: Calcium carbonate crystal morphologies of the large volume assay without protein additives. Microscopic analysis of precipitates was performed using bright field (A), fluorescence (B) and LC-PolScope in birefringent retardance (C) and orientation of the slow axis vector (D) imaging modes. The comparative analysis of the micrographs showed that no fluorescent signal was derived from precipitates without the addition of protein.

5.2.5 Modelling of Calcium-GFP Interaction

Potential calcium binding sites were predicted with the WEBfeature internet tool and illustrated with Ballview (Figure 3.13 based on the wild type GFP model no. 1GFL [Yang1996]). Acidic amino acids are concentrated in loop region I and II. This region was suggested to be responsible for calcium binding of GFP (Figure 5.5). The distances

between the α-C atoms of the selected amino acids were calculated with Ballview and were compared with the ion distances in a calcite crystal lattice (result section 3.1.4.2).

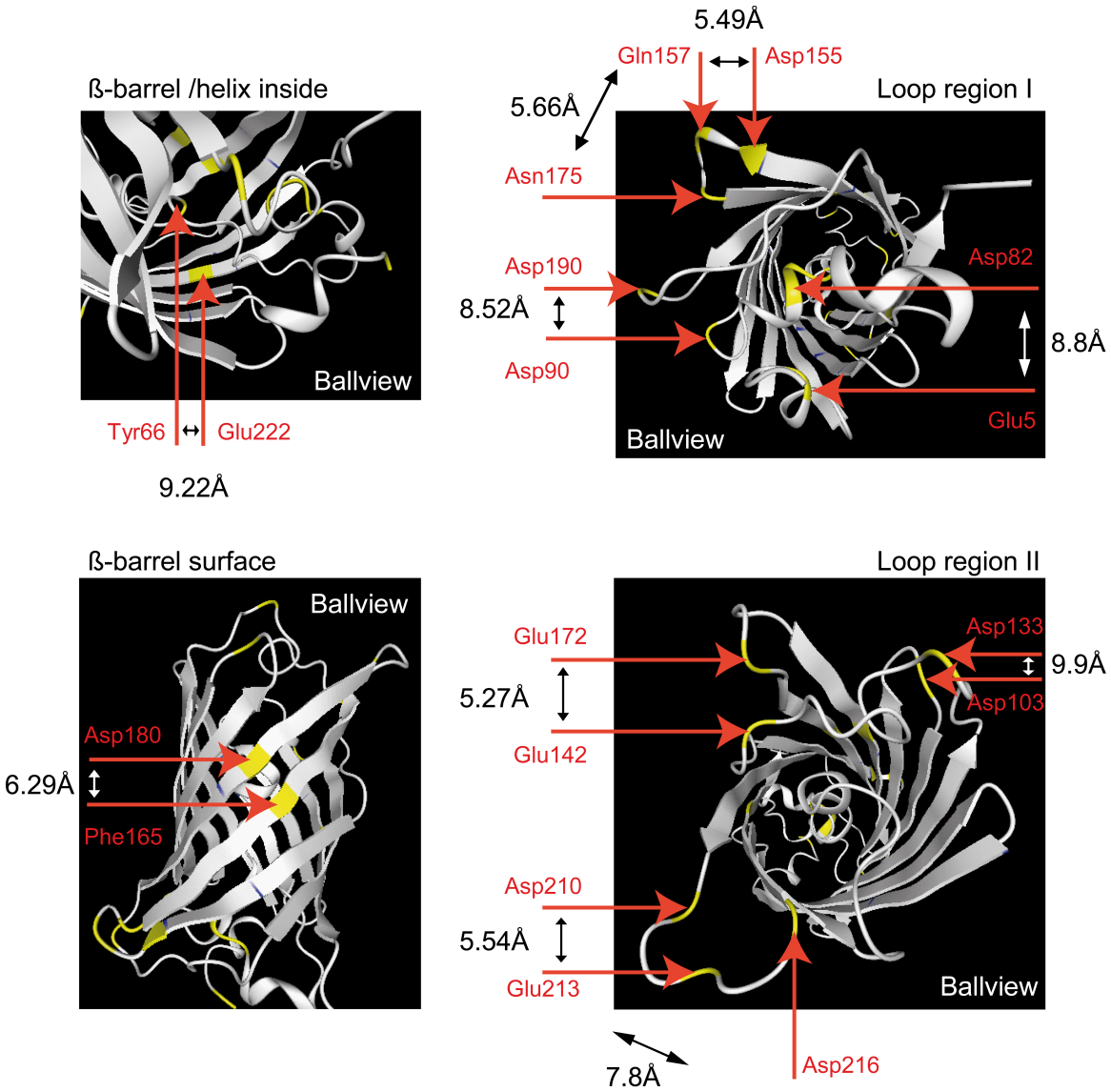


Figure 5.5: Potential calcium binding sites were predicted with the WEBfeature internet tool and illustrated with Ballview (Figure 3.13, wild type GFP model no. 1GFL [Yang1996]). Acidic amino acids are highlighted in yellow. Two potential calcium binding clusters were identified and defined as loop region I and II, each with three binding possibilities. Two additional binding sites were manually identified: Firstly, in the inside of the GFP protein between Tyr66 and Glu222 and secondly, on the surface of the β-sheet structure in between of Asp180 and Phe165. The distances between the α-C atoms of the selected amino acids are stated in Angström (Å).

Figure 5.6 illustrates a sequence alignment of amino acid compositions between wild type GFP (used for the modelling) and eGFP (used for recombinant protein expression) with clustalW. Asterisks demonstrate identical amino acids in both proteins. Dots revealed differences in the compared sequences. Double dots highlight amino acids with identical properties and single dots with different properties, compare for detailed analysis section 3.1.4.2.

```
eGFP                          MRGSHHHHHHTDPAKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDAT 50
1GFL_A|PDBID|CHAIN|SEQUENCE   ------------ASKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDAT 38
                                         .:**********************************

eGFP                          YGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAM 100
1GFL_A|PDBID|CHAIN|SEQUENCE   YGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKRHDFFKSAM 88
                              **************************::**************:********

eGFP                          PEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG 150
1GFL_A|PDBID|CHAIN|SEQUENCE   PEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG 138
                              **************************************************

eGFP                          HKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPI 200
1GFL_A|PDBID|CHAIN|SEQUENCE   HKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPI 188
                              **************************************************

eGFP                          GDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 250
1GFL_A|PDBID|CHAIN|SEQUENCE   GDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK 238
                              **************************************************

eGFP                          GGGAAKLN 258
1GFL_A|PDBID|CHAIN|SEQUENCE   --------
```

Figure 5.6: Sequence alignmnent of wild type GFP (model no. 1GFL, [Yang1996]) and eGFP used in this study [Fuhrmann1999]. Asterisks demonstrate identical amino acids in both proteins whereby dots revealed differences in the sequence. Double dots highlight amino acids with identical properties.

5.2.6 Comparison of Batch and Chromatography Purification

A comparison of GFP and GFP-perlucin eluates is shown after affinity purification with the Äkta Purifier system equipped with a His trap column (GE Healtcare) and after batch purification (Figure 5.7). The protein eluates obtained by purification with the His trap column precipitated after the elution. Therefore, only batch-purified protein eluates were used for further analysis.

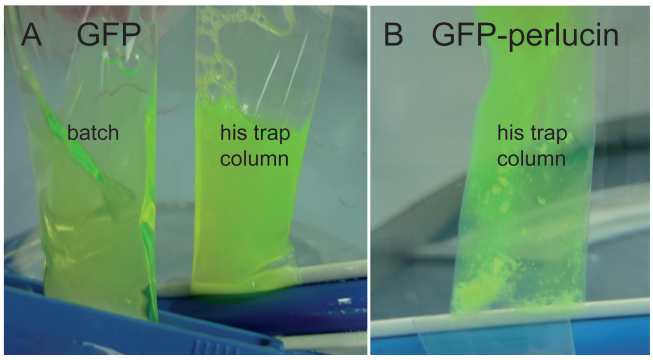


Figure 5.7: Comparison of GFP and GFP-perlucin eluates obtained after affinity purification with the Äkta Purifier system and batch purification. The GFP (A) and GFP-perlucin eluates (B) precipitated after the purification with the His trap column, exclusively. Only batch-purified protein was used for protein characterization and inorganic precipitation assays.

5.2.7 Field-Flow Fractionation - Protein Agglomeration Analysis

A schematic illustration of developed field-flow fractionation (FFF) run profiles is shown in Figure 5.8. The black line illustrates identical run conditions for each protein sample (phase I and III). The red line is part of the BSA protein (calibration standard) run profile with a constant cross flow. The blue line shows the exponential degree in the cross flow, used for GFP. The green line demonstrates a linear gradient, which was used for the GFP-perlucin protein (phase II). Detailed method parameters are described in Tables 5.1, 5.2 and 5.3.

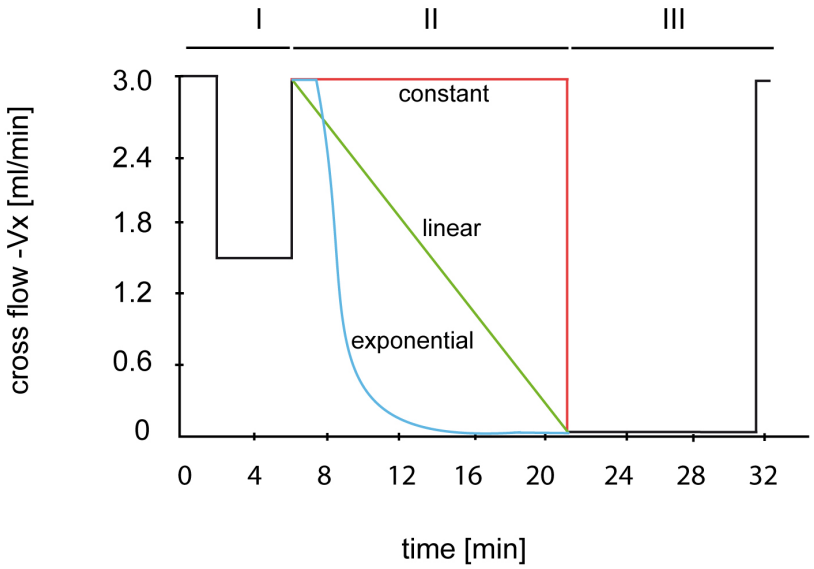


Figure 5.8: Flow profile (cross flow) for the separation of the different proteins in the channel. Phase I and III are constant for each method (black). Phase II illustrates the differences between the methods used for each protein: Red - BSA standard (constant cross flow), blue - GFP (exponential decrease in cross flow), green - GFP-perlucin (linear decrease in the cross flow). Detailed method parameter are described in Tables 5.1, 5.2 and 5.3.

Suitable concentration ranges of the proteins had to be verified empirically for a fraction separation in the field-flow fractionation. Evaluable results were obtained with BSA concentration of 1 μg/μl (inject volume of 10-30 μl), GFP concentration of 2 μg/μl (inject volume of 10-30 μl) and GFP-perlucin concentration of 6-7 μg/μl (30 μl inject volume). GFP and GFP-perlucin protein samples were measured in the precursor solutions for inorganic precipitation assays.

Table 5.1: Detailed overview of BSA method parameters.

Mode	Start	End	Duration	Vx start	Vx end
	[min]	[min]	[min]	[ml/min]	[ml/min]
Elution	0	2	2	3	3
Focus	2	3	1	0	0
Focus/Injection	3	5	2	0	0
Focus	5	6	1	3	3
Elution	6	21	15	0	0
Elution	21	26	5	0	0
Elution/Injection	26	30	2	0	0
Elution	30	31	1	3	0

Table 5.2: Detailed overview of GFP method parameters.

Mode	Start	End	Duration	Vx start	Vx end
	[min]	[min]	[min]	[ml/min]	[ml/min]
Elution	0	2	2	3	3
Focus	2	3	1	0	0
Focus/Injection	3	4	1	0	0
Focus	4	6	2	0	0
Elution	6	8	2	3	3
Elution	8	10	2	3	0.79
Elution	10	12	2	0.79	0.21
Elution	12	14	2	0.21	0.05
Elution	14	16	2	0.05	0.01
Elution	16	18	2	0	0
Elution	18	38	20	0	0
Elution	38	40	2	3	3

Table 5.3: Detailed overview of GFP-perlucin method parameters.

Mode	Start	End	Duration	Vx start	Vx end
	[min]	[min]	[min]	[ml/min]	[ml/min]
Elution	0	2	2	3	3
Focus	2	3	1	0	0
Focus/Injection	3	5	2	0	0
Focus	5	6	1	0	0
Elution	6	21	15	3	0
Elution	21	36	15	0	0
Elution	36	37	1	3	3

5.3 Expression of Perlucin-GFP in Arabidopsis thaliana

5.3.1 Genotyping of the T2 Perlucin-GFP Plants

Figure 5.4 shows the genotyping result of the t2 generation of *perlucin-GFP* transformed *Arabidopsis thaliana* plants.

Table 5.4: Gentotyping results of the t2 generation of *perlucin-GFP* transformed *Arabidopsis thaliana* plants. The EW7 (bold) line showed a reduced germination rate with only 32 % compared to the other lines and the wild type plants.

Line	Number of plants	Positives	Positives	Germination rate
			[%]	[%]
EW5	27	21	78	88
EW7	36	22	61	32
EW10	29	8	28	83
EW16	20	20	100	92
EW17	25	25	100	81
EW20	30	8	27	74
EW21	35	11	31	83
EW22	80	8	10	68
Wild type	n.a.	n.a.	n.a	88

5.3.2 RNA Extraction of T2 Plants

Quantitative real-time PCR analysis was performed to study the *perlucin-GFP* gene expression level in the different *Arabidopsis thaliana* lines. RNA was extracted from plants prior to cDNA analysis and PCR reaction. Figure 5.9 shows a representative separation of RNA on an agarose gel. The whole RNA was extracted from 12 days old *Arabidopsis thaliana* seedlings.

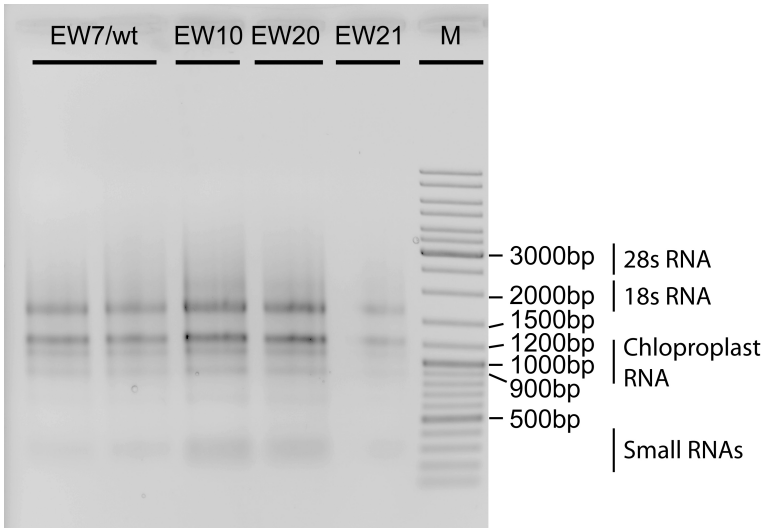


Figure 5.9: Agarose gel of extracted RNA (1-2 μg was loaded) from *perlucin-GFP* - transformed *Arabidopsis thaliana* seedlings (EW7, EW10, EW20, EW21) and wild type control (wt). Representative RNA species were obtained such as chloroplast RNA's and small RNA's. Minor amounts of 28s RNA were found. (M) Molecular weight standard, GeneRuler^{TM}.

5.3.3 Germination Rate and Root Growth Assay of T2 Generation

Figure 5.10 shows an overview of 12 days old wild type and *perlucin-GFP* transformed *Arabidopsis thaliana* plants (t2) on MS agar. The plants were used to determine the germination rate and to analyse the root growth.

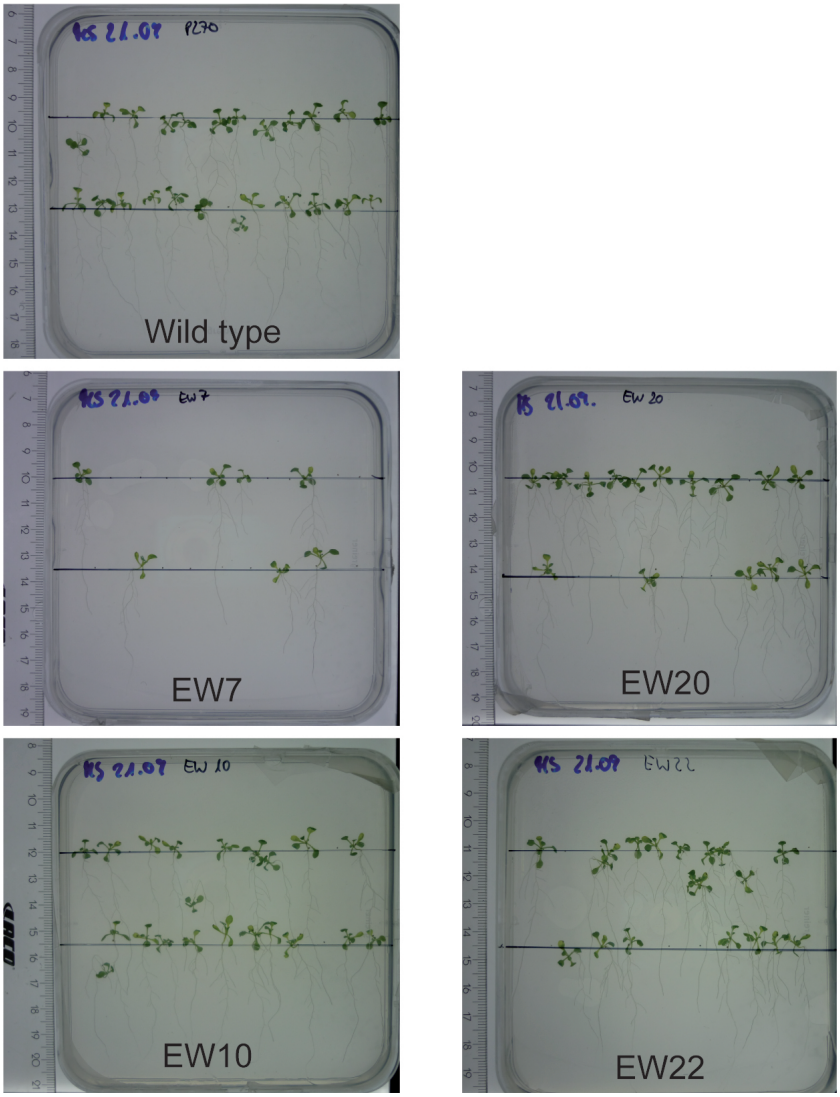


Figure 5.10: Overview of 12 days old wild type and *perlucin* transformed *Arabidopsis thaliana* plants (t2) on MS agar. The germination rate and the main root length were analysed. The EW7 line showed a reduced germination rate in contrast to wild type plants (result section, Figure 3.24). The strongest elongation of the mean root length was observed in the EW20 line, $p<0.001$ (***).

5.3.4 Genotyping T3 Plants

Detailed genotyping result of the EW7 line is shown in Table 5.5. The variation in germination rate was high among the plants (Table 5.5, bold). The plants growing on agar plates were shown in Figure 5.11.

Table 5.5: Detailed genotyping results of the EW7 line. The plants were grown on MS agar plates. The germination rate was calculated and the plants were analyzed in a PCR reaction specific for the *perlucin* gene. The variation was high between the different lines (5.1 and 24.1 in bold). See Figure 5.11 for the plants grown on agar plates.

Line EW7	Germination rate [%]	Positives in the PCR
5.1(1)	100	1
5.1(2)	100	0
5.3(1)	90	1
5.3(2)	90	0
5.4(1)	90	0
5.4(2)	95	0
5.5(1)	94	0
5.5(2)	48	0
24.1(1)	40	1
24.1(2)	47	0
24.2(1)	61	0
24.2(2)	60	0
24.3(1)	73	0
24.3(2)	65	0
24.4(1)	70	0
24.4(2)	75	0
24.5(1)	53	0
24.5.(2)	60	0

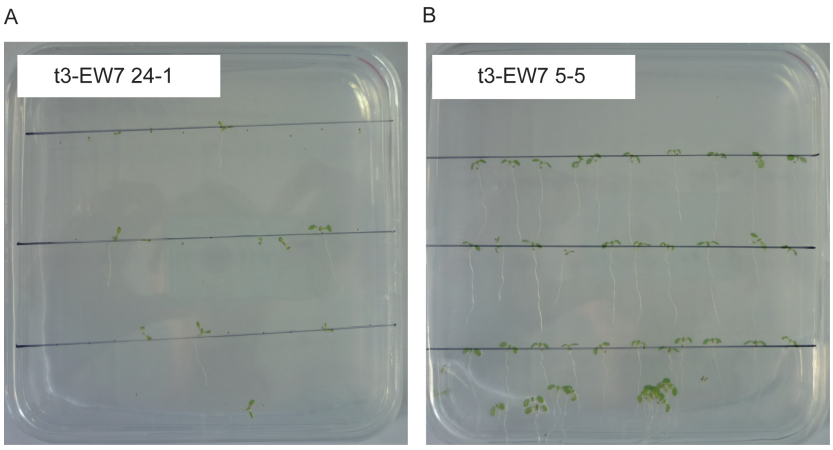


Figure 5.11: Positively *perlucin-GFP* transformed *Arabidospsis thaliana* plants of the EW7 line were obtained from plants with either 40 % (EW7-24.1, A) of germination rate or with 100 % of germination rate (EW7-5.1, B). The plants were grown on MS agar plates until genotyping.

5.3.5 Seed Analysis of T3 Plants

Absolute numbers of analyzed seeds and siliques are shown in Figure 5.12. This includes the analysis of silique length of transformed *Arabidopsis thaliana* plants compared to wild type plants. Furthermore, the absolute number of seed positions were calculated in each silique and compared to the number of empty seed positions.

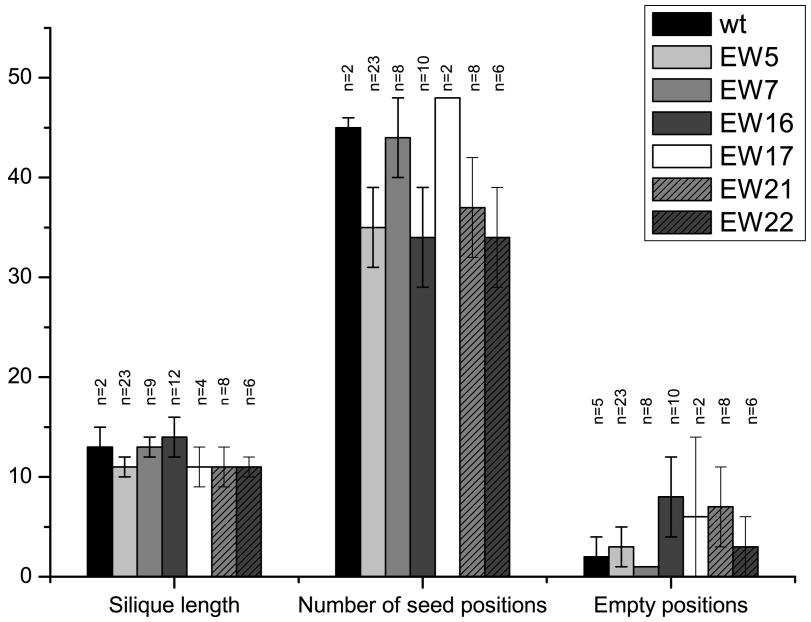


Figure 5.12: Analysis of *perlucin* transformed *Arabidopsis thaliana* plants revealed silique length, number of seed positions in the siliques and number of empty seed positions compared to wild type plants. See for relative analysis of empty seed position numbers and statistical relevance Figure 3.28. Error bars indicate the standard deviation, (n) number of analyzed siliques.

5.4 Expression of Biomineralization Proteins in Nicotiana benthamiana

The whole protein extract of transformed tobacco leaf cells was analyzed with an anti GFP Western blot procedure. The insoluble fraction after extraction is shown in Figure 5.13. The analysis of soluble protein fraction is shown in the result section 3.2.5.2.

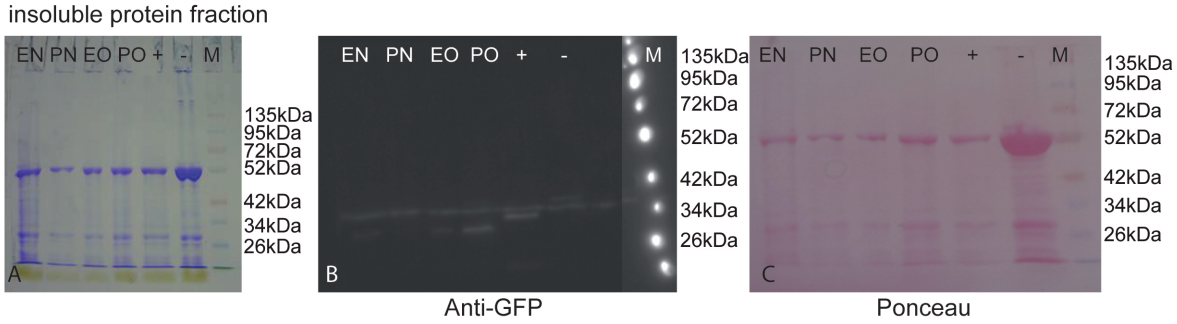


Figure 5.13: Insoluble protein extract from transformed tobacco leaves (after 48h) were separated on a SDS-PAGE (A). OC17-GFP (46 kDa) and N16N-GFP (34 kDa) were detected with an anti GFP antibody (HRP, B). All samples bore a band (less than 34 kDa), which represented a GFP degradation product. The GFP positive control (29.7 kDa) migrated on a size slightly smaller than 34 kDa. The successful transfer from the protein to the membrane was shown with a Ponceau staining (C). (EN) EXT3-N16N-GFP, (PN) PRP4-N16N-GFP, (EO) EXT3-OC17-GFP, (PO) PRP4-OC17-GFP, (+) GFP, (-) negative control, (M) molecular weight marker, Spectra^{BR}.

Publications

Parts of this work were published in:

Weber, E., Guth, C., Eder, M., Bauer, P., Arzt, E., Weiss, I. M., Biotechnological Mineral Composites via Vaterite Precursors, Mater. Res. Soc. Symp. Proc. Vol. 1465, DOI: 10.1557/opl.2012.1046

Weber, E., Guth, C., Weiss, I.M., GFP Facilitates Native Purification of Recombinant Perlucin Derivatives and Delays the Precipitation of Calcium Carbonate, PLoS ONE 7(10): e46653. doi:10.1371/journal.pone.0046653

Weiss, I.M., Weber, E., Eder, M., Arzt, E., Schneider, A.S.,Neue organisch-anorganische Kompositmaterialien durch Biomineralisation, DE102011057183.3, 30.12.2011 (Priorität)