The Cytoskeleton of Diatoms

Structural and Genomic Analysis

Dissertation

zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Charlotte Aumeier
aus
Nürnberg

Bonn, Oktober 2014

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

Erstgutachter:	Prof. Dr. Diedrik Menzel
Zweitgutachter:	Prof. Dr. Peter Kroth
Tag der Promotion:	12.12.2014
Erscheinungsjahr:	2015

Eidesstattliche Erklärung

Hiermit versichere ich, dass diese Dissertation von mir selbst und ohne unerlaubte Hilfe angefertigt wurde. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Ferner erkläre ich, dass die vorliegende Arbeit an keiner anderen Universität als Dissertation eingereicht wurde.

Bonn, 01.Oktober 2014

_____

Charlotte Aumeier

Preface

Ever heard of diatoms, one of the largest and ecologically most significant groups of photosynthetic unicellular organisms on earth? Sure you have, and if it is only for their beautifully ornamented, silicified cell wall.

The cytoskeleton of diatoms intrigued cell biologist for over a century now. The present study carries on this passion, with the aim to further contribute towards the understanding of structure and function of the actin and microtubule system in these fascinating organisms. For this intention a variety of available methods have been used, some others have been tested but found inadequate and yet others had to be optimized. Starting off with classical cell biological methods like transmission electron microscopy of resin embedded cells and laser scanning confocal microscopy of samples treated with fluorescent probes and after immunofluorescence labeling. It goes right on with genetic transformation approaches to visualize aspects of the cytoskeleton *in vivo* by the expression of reporter constructs, and continues all the way to an extensive *in silico* analysis of genomic sequence data from publicly available databases with focus on actin, actin-related proteins (ARPs) and actin-binding proteins (ABPs).

The first chapter deals with diatom culturing techniques for laboratory use. Further, the results of the indirect immunofluorescence labeling of actin and tubulin in *Craticula cuspidata* are presented in the second chapter. A new transformation protocol is described in chapter three, which did not lead to any transformants for *C. cuspidata* but was successful in *Cylindrotheca fusiformis* using Lifeact-GFP as an actin reporter construct. Finally, chapter four presents the *in silico* analysis of the inventory of actin, ARPs and ABPs in the diatoms.

Contents

Culturing Diatoms

1 Introduction

Diatoms are a divers group of algae, varying in their requirements, demanding different media, and different handling techniques. The species used in the present study for the study of cytoskeletal organization, *Phaeodactylum tricornutum*, *Cylindrotheca fusiformis* and *Craticula cuspidata*, are all members of the pennate diatoms. *C. fusiformis* and *P. tricornutum* are salt water diatoms, whereas *C. cuspidata* is a fresh water diatom. *C. cuspidata* cells are strongly adherent to the substratum, whereas *C. fusiformis* and *P. tricornutum* preferentially grow in suspension. The genome of *P. tricornutum*, the most commonly used diatom model organism, is sequenced (Bowler et al., 2008) and genetic transformation protocols are established (Apt et al., 1996; Zhang and Hu, 2013). A genetic transformation protocol was also established for *C. fusiformis* (Fischer et al., 1999), but the species was less frequently used for genetic studies. In contrast, no sequence data or transformation methods are available for *C. cuspidata*, since it was rarely a subject of studies in the past.

Over the last 20 years the technical possibilities for doing research on diatoms has advanced considerably, with established genetic transformation methods as well as full genome sequence informations available for some diatoms such as *Thalassiosira pseudonana* and *P. tricornutum*. All of the commonly used diatom model organisms, for which transformation protocols are available, are relatively small in size. *P. tricornutum* for example, has a cell diameter of about 2 μm, which is too small for any detailed microscopic analysis of internal structures. Therefore *C. cuspidata* was adopted as an alternative, because of its size (a length of about 100 μm, a width of about 20 μm), and also because of its rather short cell cycle of approximately one division per day.

In the following, the setting up of laboratory cultures for *P. tricornutum* as well as *C. fusiformis* is described, and basic cell culture parameters to grow *C. cuspidata* under laboratory conditions are investigated.

2 Material & Methods

2.1 Strains

Cell biological work and genomic manipulation was performed with the three diatom species: *C. cuspidata*, *P. tricornutum* and *C. fusiformis* (Tab. 2.1).

Table 2.1: Diatom strains used in this study. *C. cuspidata* was isolated by Sebastian Hess and is now available from the algal culture collection (CCAC) maintained by Dr. Barbara Melkonian at the Cologne Biocenter.

Strain	Genotype	Reference
Craticula cuspidata	Wilde-type	collected by S. Hess
Cylindrotheca fusiformis	Wilde-type	collection of N. Poulsen
Phaeodactylum tricornutum	Wilde-type (Pt4)	collection of P. Kroth

2.2 Culture media

C. cuspidata was grown in Waris-H (previously described by McFadden and Melkonian, 1986) with additional 3 x vitamins and 0.50 mM Na_2SiO_3 x $9H_2O$, but without soil extract. *C. fusiformis* was cultured in artificial see water (previously described by Poulsen and Kröger, 2005) and *P. tricornutum* was cultured in f/2 medium (previously described by Guillard and Ryther, 1962). For plate cultures, 1,2% agar was added to the respective medium, and when required appropriate antibiotics were also added (Tab. 2.2).

Table 2.2: Antibiotics used in this study.

Antibiotic	Stock	Final	Dissolvent
Nourseothricin (ClonNat)	2 mg/ml	0.2 -100 mg/l	ddH_2O
Geneticin (G418)	10 mg/ml	1-100 mg/l	ddH_2O
Zeocin	100 mg/ml	1 - 1000 mg/l	ddH_2O
Formula I (James, 1974)	0.6 g Penicillin 'G', 1 g Streptomycin Sulfate	1 ml/100 ml	200 ml ddH_2O

2.3 Diatom cultivation

C. fusiformis and *P. tricornutum* were grown under aerobic conditions at 17 °C in liquid culture medium with a day to night cycle of 16 h:8 h. The light intensity was 35 μE $m^{-2}s^{-1}$ and warm white light from Osram fluorescent tubes (Lumilux L36W/830) was used. The liquid culture batches (150 ml) were shaken at 150 rpm.

Basic culturing conditions were investigated and optimized for *C. cuspidata*. To determine the optimal illumination for *C. cuspidata* the light intensity was varied in steps of 15 μE $m^{-2}s^{-1}$ between 20 μE $m^{-2}s^{-1}$and 80 μE $m^{-2}s^{-1}$. The day and night cycle was varied between 8 h:16 h, 12 h:12 h; 16 h:8 h and continuous light. To optimize the culturing temperature, settings of 8 °C, 12 °C, 17 °C, 20 °C, and 24 °C were tested. The division rate in liquid cultures was investigated, as well as growth rate and colony formation on solid agar plates, required to select transformants after microprojectile bombardment (chapter 10.5.1).

To evaluate the effects of cultivation parameters on the viability of cultured cells, microscopic appearance of the cells was rated on a scale from 1 to 5, with 1 for cell at excellent state and 5 for dead cells: 1 — no vacuolization; 2 — little vacuolization around the nucleus; 3 — vacuoles around the nucleus; 4 — highly vacuolated cell with displaced plastid; 5 — green, displaced plastids, the cells were dead.

2.4 Axenic *C. cuspidata* culture

To obtain an axenic culture of *C. cuspidata*, different isolation and purification techniques were tested, such as: micropipette washing technique; atomizer technique; washing by centrifugation technique; antibiotics technique; sonification washing technique and the filter technique described in detail by James (1974). None of these attempts yielded satisfactory results.

C. cuspidata produces large amounts of mucilage, which is interspersed with bacteria and empty frustules. Therefore a new purification method was developed, combining rotation on a horizontal shaker, followed by rapid centrifugation procedures. For this procedure the adherent *C. cuspidata* cells were scraped off the culture dish and resuspended in fresh culture medium. This suspension was slowly rotated on a shaker, which caused the mucilage to aggregate at the border of the culture dish, whereas the diatoms accumulated in the center. The diatoms were harvested and transfered into a 50 ml Falcon® tube and resuspended in fresh medium. Centrifugation at 2000 g for 1 min separated the diatoms (pellet) from the bacteria (supernatant). The supernatant was discarded, the pellet resuspended in fresh medium and the centrifugation step repeated four times. 500 ml fresh culture medium was inoculated with 200 μl washed *C. cuspidata* cells.

2.5 Antibiotic treatment to select transformed *C. cuspidata* after particle bombardment

To determine, which selection antibiotic would be best suited to identify transformants carrying the antibiotic resistance gene, an antibiotic had to be chosen that reliably kills untransformed cells of *C. cuspidata*. Three antibiotics, Zeocin, ClonNat, and G418, were tested. Survival rate and morphological appearance of cells grown on Waris-agar plates containing the different antibiotics were monitored, with at least three replicates and compared with cells grown under the same conditions, but without antibiotics. A broad concentration spectrum for each antibiotic was tested and observed over a period of one month: 1 μg/ml, 5 μg/ml, 15 μg/ml, 25 μg/ml, 50 μg/ml, 75 μg/ml, 100 μg/ml. To evaluate the toxicity of the antibiotic towards *C. cuspidata*, the morphology of the cell was determined and classified from state 1 to state 5, with state 1 describing excellent growth and state 5 dead cells (for detail see 2.3).

3 Results

3.1 Diatom cultivation

C. cuspidata was chosen in this study as a model organism for genetic manipulation. Since all protocols for isolation of diatom transformants so far rely on antibiotic selection of colonies on solid media, growth on agar plates and building colonies is hence a basic prerequisite.

C. cuspidata tolerated temperatures from 8 °C to 24 °C. Good and stable growth conditions with a cell cycle of 24 h to 48 h were achieved at 17 °C, so that new culture batches needed to be set up only every fifth week. *C. cuspidata* grew and proliferated well on solid agar plates and was capable to form colonies under these conditions (Fig. 3.1), although better cell conditions and a higher division rate were achieved in liquid medium. Mucilage production is a constant problem in laboratory cultures of diatoms, because it promotes bacterial growth. Cells grown in plastic petri dishes tended to produce less mucilage, in comparison to those grown in glass ware. Additional aeration of the liquid culture batch lead to even more mucilage secretion, and therefore had to be avoided.

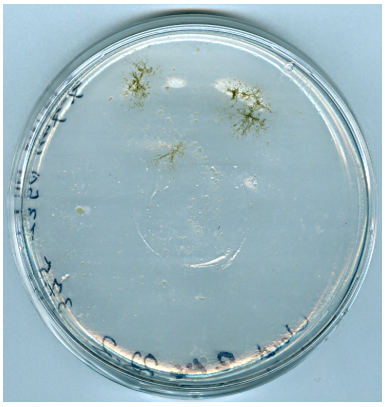


Figure 3.1: *C. cuspidata* growing in colonies on agar plates.

Continuous illumination for more than seven days was lethal for the culture and *C. cuspidata* did not survive high light conditions of 75 μE $m^{-2}s^{-1}$ at a 16 h:8 h cycle over a longer period.

Good culturing parameters meeting the needs of all three organisms used in this study, *C. cuspidata*, *P. tricornutum* and *C. fusiformis* were a 16 h:8 h light:dark cycle with 35 μE $m^{-2}s^{-1}$ at 17 °C,

3.2 Axenic *C. cuspidata* culture

To purify and isolate *C. cuspidata*, different techniques were verified. The only method that resulted in axenic cultures was the antibiotics technique described by James (1974). Using the antibiotic mixture, Formular I of this author, reduced the number of bacteria efficiently within 48 h, while *C. cuspidata* survived well on the Formular I-containing agar plate for over 5 days. An unexpected problem occurred, when the harvested bacteria-free cells were transferred into fresh, sterile Waris medium. None of the axenic cultures survived, as all cells died within a period of 8 days.

Therefore a two step purification method was developed in this study, separating diatoms from the mucilage (interspersed with empty frustules) and the bacteria growing on them by centrifugal force, followed by a rapid centrifugation reduced the contamination, although it did not lead to an axenic culture. The mechanical separation could not remove all bacteria, since some bacteria kept sticking to the frustules of the cells, but these few bacteria did not interfere with the diatom growth over several weeks. Cells of the freshly inoculated culture batches divided faster and showed a good

morphology with low lipid production and little vacuolization, when compared to contaminated cultures.

3.3 Antibiotic treatment to select transformed *C. cuspidata* after particle bombardment

The broad-spectrum antibiotic Zeocin from *Streptomyces verticillus*, frequently used for antibiotic selection of transformants, had no effect on *C. cuspidata*. Cells grown on plates containing 100 μg/ml Zeocin for one week were comparable in their appearance and viability to control cells. Therefore, Zeocin was not chosen for the selection of transformants.

The antibiotic G418 (geneticin), revealed a severe toxic effect after three days on *C. cuspidata*. 25 μg/ml G418 was lethal to all cells within one week, whereas they survived on 1 μg/ml G418 plates for over a month. In a further concentration series, it was determined that a concentration of 20 μg/ml G418 over a period of 6 days is as effective as 25 μg/ml (Fig.3.2), but that within this time 25 μg/ml G418 additionally kills all bacteria. The streptothricin antibiotic ClonNat, which has been developed as an alternative selection antibiotic for the screening of recombinant cells (Krügel et al., 1993), had an even stronger effect towards *C. cuspidata*. 5 μg/ml of ClonNat was already lethal for most cells after three days and even at a concentration of 1 μg/ml, no living cells were observed after one month. A concentration of 2 μg/ml ClonNat was efficient to kill all cells within six days (Fig.3.2). Therefore, these two antibiotics at the determined threshold concentrations (25 μg/ml G418 and 2 μg/ml ClonNat) were chosen for the selection screens employed after genetic transformation (chapter 10.5.1).

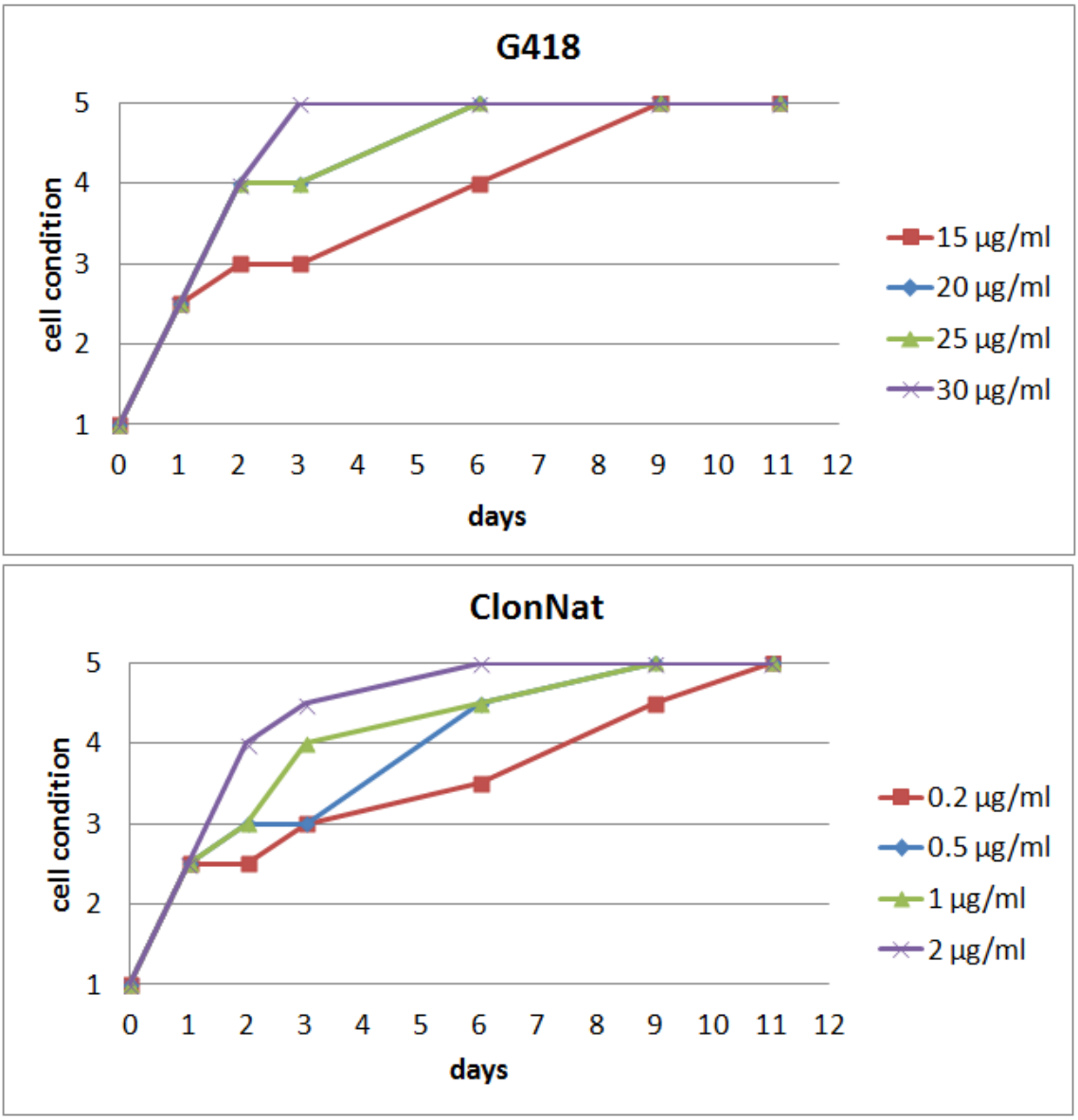


Figure 3.2: Response of *C. cuspidata* to different concentrations of G418 and ClonNat. The toxic effects on cell morphology of *C. cuspidata* exerted by 15-30 μg/ml G418 and 0.2-2 μg/ml ClonNat are plotted against time. Criteria for the classification of cell states are: state 1 — normal appearance/no vacuolization; state 2 — little vacuolization around nucleus; state 3 — vacuoles around the nucleus; state 4 — highly vacuolated cell with displaced plastid; state 5 — plastids are green and displaced, cells are dead. Control cells settle between states 1 and 2.

4 Discussion

4.1 Diatom cultivation

Settings for culture maintenance and reproduction under laboratory conditions of the diatom species *P. tricornutum*, *C. fusiformis* and *C. cuspidata* were established as a prerequisite for the molecular genetic and cell biological experiments described in the present study. Optimization of culturing parameters for *C. cuspidata* lead to a cell cycle length of 24-48 h, which is a reasonable time span to allow colony formation on agar plate cultures within 2-3 weeks after genetic transformation.

4.2 Axenic *C. cuspidata* culture

Diatom cells isolated from their natural habitat invariable have bacteria attached to the frustule, or associated with the mucilage which adheres to the cells. While some bacteria might be beneficial for sustained growth, others may just feed on mucilage and other secreted metabolites. When cultivated under laboratory conditions, bacterial contamination will eventually befoul the resources in the narrow confines of a culture dish. It is therefore important to reduce or even completely remove this bacterial contamination. In oder to establish an axenic culture of *C. cuspidata* different purification methods have been tested. Treating the culture with the antibiotic mixture Formula I (James, 1974) selectively was lethal for bacteria, while it did not visibly interfere with the cell condition of *C. cuspidata*. However, this treatment did not lead to a stable axenic culture, because all cultures that were started with the axenic *C. cuspidata* cells, died after several days.

The Formula I antibiotic mixture possibly had a delayed toxic effect on the cells. So that they survived for 5 days on the antibiotic plate, but were affected to such an extent by the antibiotics, that they died afterwards. Another possibility could be that *C. cuspidata* needs the interaction with certain bacteria to sustain growth. This hypothesis is difficult to ascertain for this particular species, due to the tight association of the cells with bacteria. However, there is experimental evidence showing that the growth of diatoms in general can be enhanced or suppressed by the presence of bacteria (Baker and Herson, 1978; Grossart et al., 2006; Grossart and Simon, 2007; Bruckner et al., 2008; Amin et al., 2012).

Other purification methods, which were tested in the current study, such as micropipette washing technique, atomizer technique, sonification washing technique, failed due to the extensive mucilage secretion by *C. cuspidata*. The combined rotation and centrifugation technique developed in this study, has the advantage that most mucilage interspersed with empty frustules and bacteria got separated from the living cells due to their different densities. Carefully adjusted speed of rotation lead to the effect that the aggregated mucilage interspersed with empty frustules was pulled outwards by centrifugal forces, while the living cells concentrated in the middle of the culture dish by an opposing force. This phenomenon is further supported by the adhesion of the living cells to the culture dish. After the cells were separated from most of the mucilage the diatoms were

centrifuged at low speed. The centrifugation separated bacteria from the diatoms, at a low speed the heavy diatoms sank to the bottom of the tube, whereas the light bacteria remained in the supernatant. This significantly reduced the amount of bacteria in the suspension. Repetition of centrifugation after resuspension of the pellet further cleaned the cells, but never resulted in an axenic culture. This was most probably caused by the fact, that some bacteria stuck so tightly to the frustules, that they could not be removed. The mild mechanical separation of cells and mucilage/bacteria via rotation and centrifugation did not harm the cells of *C. cuspidata*, as they retained their morphology and high division rate over several weeks.

In summary, bacterial contamination does become a problem, when their number passes a certain threshold. This may be caused by toxic secretions of the bacteria or by the limitation of nutrients as has been suggested previously by other authors (Baker and Herson, 1978; Grossart et al., 2006; Grossart and Simon, 2007; Bruckner et al., 2008). Therefore, the routine strategy for culturing *C. cuspidata* was to reduce bacterial contamination every third month by mechanical separation via rotation and centrifugation.

4.3 Antibiotic treatment to select transformed *C. cuspidata* after particle bombardment

To select *C. cuspidata* transformants after genetic manipulation, antibiotics that have been used as selection agents for diatoms in previous studies, Zeocin, ClonNat and G418 (Zaslavskaia and Lippmeier, 2000), were tested. Zeocin is a broad spectrum antibiotic, which is thought to act by intercalating and cleaving DNA (Drocourt et al., 1990), whereas ClonNat and G418 inhibit the protein biosynthesis (Bar-Nun et al., 1983; Krügel et al., 1993; Kochupurakkal and Iglehart, 2013). The results presented here, show that the antibiotics inhibiting protein biosynthesis, are more toxic for *C. cuspidata* than Zeocin, leading to cell death at concentrations as low as 1 μg/ml, whereas Zeocin was ineffective even at concentrations of up to 100 μg/ml. For comparison the published antibiotic concentration to selection *C. fusiformis* transformants is 1000 μg/ml Zeocin (Poulsen and Kröger, 2005). The effect of Zeocin at higher concentrations was not further investigated, as two antibiotics (ClonNat and G418) suitable for selection were found.

A convenient by-effect of G418 is, that it is lethal for bacteria. This inhibits the competitive growth of bacteria during selection of transformants. On the other hand it cannot be excluded, that *C. cuspidata* only survives when certain bacteria are co-cultured. Therefore both, ClonNat and G418 have been used further on for the selection of transformants.

Analysis of cytoskeletal organization in the pennate diatom *C. cuspidata*

5 Introduction

Diatoms bear a great potential for various nanotechnological applications and the attention they attract from diverse scientific fields is growing: i.e., material science, (bio)chemistry, computer science and last but not least biology. Diatom cells are enclosed within a unique, beautifully sculptured, siliceous wall, called the frustule. These frustules are extraordinarily complex in symmetry and in the diversity of morphological features, which are the basis of the taxonomic description of the over 100,000 diatom species. The frustule elements, such as valves and girdle bands, are prefabricated in large cisternal structures termed the silica depositing vesicles (SDVs). Only after completion they are secreted by a huge exocytosis process to become inserted into the frustule framework on the outside of the cell (reviewed in Aumeier and Menzel, 2012). The complex nano-scale frustule patterning occurs within the SDV, but there is a meso- to micro-scale shaping that also occurs due to the forces exerted on the SDV membrane by the intracellular cytoskeleton and organelles (e.g. Mitochondria) (Pickett-Heaps and Kowalski, 1981; Schmid, 1987; Pickett-Heaps et al., 1990; Schmid, 1994; Van de Meene and Pickett-Heaps, 2002, 2004; Kröger and Poulsen, 2008; Tesson and Hildebrand, 2010a,b). The microtubule (MT) and actin cytoskeleton are tightly associated with the SDV membrane, and cytoskeletal inhibitors have been shown to influence the diatom valve morphogenesis, indicating their functional importance in this process (Cohn et al., 1989; Schmid, 1984).

The diatoms can be divided by their frustule morphology into two lineages the centrics and the pennates. Among the pennates are the raphid pennates, a monophyletic class defined by the possession of an elongated slit through the valve, the raphe. The raphe, is a requirement for diatom gliding, or locomotion. Drum (1963) discovered two thick actin bundles in the cortical cytoplasm parallel to the raphe and Edgar and Pickett-Heaps (1984) proposed an actin-based model of pennate diatom gliding. This model involves secreted adhesive fibers firmly attached to the substratum, while they remain connected to the raphe actin bundles. Poulsen et al. (1999) extended the model by including myosins as the motor elements, bridging between the adhesive fibers and one of the two prominent actin bundles that run parallel to the raphe. Since myosins remain connected with the adhesive fibers, the force generated by them leads to a backward translocation of the whole cell, relative to the stationary attached myosin-adhesive complex. This complex only detaches and falls off, when it reaches the end of the bundle.

Little is known about the assembly and molecular architecture of the two actin bundles. It is still unclear, where the filaments of the bundles nucleate, whether they are oriented in a parallel, anti-parallel, or mixed polarity fashion, and how they are anchored to the plasma membrane (Poulsen et al., 1999).

The more recent studies of the diatom cytoskeleton are concerned with the valve morphogenesis or locomotion, whereas in early studies beginning already at the end of the 19th century scientists had investigated the cell devision. Fascinating pioneer research of early scientists such as Robert Lauterborn on mitosis and cytokinesis in selected diatom species inspired great scientists in more

recent times. Such as Jeremy Pickett-Heaps, and Zac Cande generating the highest-quality optical and EM descriptions of nuclear division, spindle development and chromosome segregation so far available for diatoms (reviewed in De Martino et al., 2009). Some of these articles deal with the combined application of MT immunofluorescence labeling, with pharmacological treatments using agents directed against MTs (Wordeman et al., 1986; Wordeman and Cande, 1990; Spurck and Pickett-Heaps, 1994). However, except for the ultrastructural work, there is little publication on the MT cytoskeleton in the pennates on the level of immunofluorescence labeling in combination with confocal microscopy.

The actin cytoskeleton is not studied as extensively as the MT cytoskeleton in diatoms. For the visualization of the actin cytoskeleton, fluorochrome-conjugated phalloidin has been used as a specific probe combined with actin inhibitors. This research, also including detailed ultrastructural analysis, led to the discovery of the two raphe actin bundles in the pennates and the ring like structures at the base of the proboscis in the centric *Proboscia* and *Rhizisolenia*, which turn out to be equivalent to the actin ring at the leading edge of the SDVs involved in valve morphogenesis (Poulsen et al., 1999; Van de Meene and Pickett-Heaps, 2002, 2004; Tesson and Hildebrand, 2010a). Only the publication of Tesson and Hildebrand (2010b) revealed some network-like structures at the SDV, apart form the bundled structures, but the images have not been captured by confocal microscopy and are therefore not highly detailed.

Regarding the fundamental role of the cytoskeleton and the upcoming trend to use diatoms as model organisms, there is an increasing need to further our knowledge about the structural organization of the MT and actin cytoskeleton within diatoms has to be increased. This gap in our knowledge, concerning both the actin and the MT cytoskeleton, is filled by the present study using the pennate diatom *Craticula cuspidata* as the study subject.

6 Material & Methods

6.1 Drug treatments

C. cuspidata was treated with the actin drug jasplakinolide at a concentration of 1 μM and the MT depolymerizing drug oryzalin at a concentration of 0.5 μM and 1 μM. Cells were fixed immediately after treatment.

6.2 Indirect immunofluorescence labeling of actin and tubulin

Cells were collected by sedimentation and fixed for 15 min with a solution of 0.5 % glutaraldehyde and 1.5 % formaldehyde in actin stabilizing buffer (ASB: 50 mM Pipes, 10 mM EGTA, 5 mM $MgCl_2$, 50 mM KCl, 1 % DMSO, pH 7.1) under constant movement on a rocking tablet. Cells were rinsed twice with ASB and placed in a small filter with a mesh size of 10 μm. Between each of the following steps the cells were rinsed three times for 5 min under constant rocking. Cells were treated with fresh $NaBO_4$ in ASB twice for 15 min and afterwards lysed twice for 30 min under constant rocking using a permeabilisation solution (50 mM Glycine, 1 % Triton X-100, 1 % BSA, 1 % gelatin fish in ASB). To reduce background staining, the cells were treated with 1 % BSA, 0.1 % fish gelatin in PBS (0.14 M NaCl, 2.7 mM KCl, 6.5 mM Na_2HPO_4*$2H_2O$, 1.5 mM KH_2PO_4, pH 7.5) for 30 min under constant rocking. After rinsing in PBS a 1:100 dilution of the appropriate actin, or tubulin antibody in PBS (1 % BSA) was added and incubated at 37 °C for 2 h and then left at RT over night. The cells were then rinsed with PBS and labeled with the secondary antibody at a 1:100 dilution in PBS (1 % BSA) at 37 °C for 2 h. The cells were rinsed, counter-stained with a 1:2000 dilution of 4',6-diamidino-2-phenylindole (DAPI) in PBS for 10 min, mounted on a microscope slide using antibleach (0,1 % [w/v] para-phenylenediamine prepared in 90% [v/v] glycerol in PBS) and observed by confocal microscopy.

The following primary antibodies were used: mouse monoclonal anti-actin N350 (IgM, Amersham Pharmacia), polyclonal anti-actin MA1-744 (IgG, Thermo Scientific), anti-α-Tubulin TUB-1A2 (IgG, Sigma), polyclonal anti-γ-Tubulin (IgG, abcam). Secondary antibodies were anti-mouse (IgM), anti-mouse (IgG), anti-rabbit (IgG) conjugated with Alexa Fluor 488 or 546, respectively (Invitrogen).

6.3 Microscopy

For light microscopy the cells were immobilized on 1 % polyethylenimine-(PEI)-coated slides and examined with a Leica DMRBE conventional epifluorescence microscope.

Confocal imaging was performed using an Olympus FV1000 confocal laser scanning microscope using a 60x/1.35 (oil immersion) UPlanSApo objectives (Olympus), operated with the FV 10- ASW 3.0 software (Olympus, Hamburg, Germany). Alexa488-conjugated antibodies were excited at 488 nm and detected between 500-550 nm. Alexa546-conjugated antibody were excited at 543 nm and

detected at 550-650 nm. The fluorescent dye DAPI was excited at 405 nm and detected at 420-470 nm and the autofluorescence of the plastid was excited at 635 nm and detected at 655-755 nm. To avoid bleed-through in double labeling experiments, all images were captured using the sequential line scanning mode. Image processing was done with the Olympus software FV1000 (Ver.3.0), GIMP 2.8, and Open Office applications.

6.4 TEM

Cells of *C. cuspidata* were collected by sedimentation and fixed for 30 sec with a solution of 0.5 % glutaraldehyde in Waris fix (1.00 mM KNO_3; 81.10 μM $MgSO_4$ x 7 H_2O; 0.15 mM $(NH_4)_2HPO_4$; 0.42 mM $Ca(NO_3)_2$ x 4 H_2O; 1.00 mM HEPES; 1 mM EGTA; 1 % DMSO; pH 7.1). After 30 sec osmium tetroxide was added to a final concentration of 1 %. After 10 min at room temperature the solution was exchanged with 1 % osmium in ddH_2O and further fixed for 30 min. The washed cells were dehydrated slowly in a graded series of acetone and embedded in Spurr's resin (Komis et al., 2002). Ultrathin sections cut on an ultracut II (Reichert), were contrasted with uranyl acetate and lead citrate, and examined with a LEO 912AB electron microscope.

7 Results

The cytoskeleton of *C. cuspidata*, a motile pennate diatom with a length of about 100 μm and a width of about 20 μm, was visualized by indirect immunofluorescence labeling. The cell structure is presented in valve view, looking from the top down to the bottom of the cell, as well as in girdle band view, looking through the cell from one side to the other, by confocal microscopy. The focus of the present study is placed on the mature interphase stage, when cell morphogenesis is completed. Therefore, the actin ring at the advancing edge of the SDVs reported in previous publications (Poulsen et al., 1999; Van de Meene and Pickett-Heaps, 2002; Tesson and Hildebrand, 2010a), could be confirmed, but was only occasionally observed (data not shown).

7.1 Visualization of the actin cytoskeleton

The raphe (Fig. 7.1 E) spanned the cell from one cell pole to the other, but was interrupted in the middle by the central nodule (Fig. 7.1 F). The two characteristic actin bundles (Fig. 7.1 B, D) aligned parallel underneath the raphe, spanning the full cell length. The TEM images confirm, that the two raphe actin bundles were tightly associated with the plasma membrane located just under the raphe (Fig. 7.1 E). The girdle band view of the cells clearly shows, that the two raphe actin bundles of the hypotheca were not connected with the two raphe actin bundles of the epitheca (Fig. 7.1 C).

Confocal microscopy revealed divers and complex actin structures that were best appreciated in single optical sections. Apart from the thick raphe actin bundles *C. cuspidata* also had an extensive actin network spanning the cell. Just underneath the plasma membrane, the area correlating with the cell cortex, had a fine actin meshwork including rather straight, long filaments in a predominantly radial organized array, protruded from the central nodule outwards to the cell periphery (Fig. 7.1 D1). Optical sections captured further inwards of the cell show an extensive actin network consisting of actin filaments, and slightly bundled filaments, without any obvious favored orientation (Fig. 7.1 D2). This major cytoplasmic network spanned the whole cell with the highest density around the nucleus.

7.2 The effect of jasplakinolide on the actin cytoskeleton

Jasplakinolide is an actin drug, promoting actin polymerization and stabilizes actin filaments. *In vivo* this causes a disruption of the actin cytoskeleton, leading to an aggregation of actin fragments into larger clusters also referred to as paracrystalline aggregates (Holzinger, 2010; Lázaro-Diéguez et al., 2008). This inhibitor is used in the current study for the first time on diatoms. An alternative agent, latrunculin B, frequently used to study actin-dependent processes in higher and lower plants as well as protists (Gibbon et al., 1999; Poulsen et al., 1999; Morton et al., 2000; Yarmola et al., 2000), had only weak inhibitory effects on *C. cuspidata*. In *C. cuspidata* both latrunculin B and its stereoisomer, latrunculin A, have been tested here at concentrations from 1 μM to 5 μM, for

up to 24 h. Neither locomotion was impaired nor did any change occur in the structure of the actin cytoskeleton. This is in marked contrast to the effects of jasplakinolide, which changed the organization of the actin cytoskeleton in *C. cuspidata* tremendously after a 24 h treatment at a concentration of 1 μM.

Upon the 1 h jasplakinolide treatment, the raphe actin bundles and the major cytoplasmic network rearranged into shorter bundled F-actin, orientated predominantly along the longitudinal axis of the cell (Fig. 7.2 A). The polymerization of new actin filaments seemed to occur along the raphe and around the central nodule. The fine cortical actin meshwork was still visible in the cell periphery after 1 h jasplakinolide treatment.

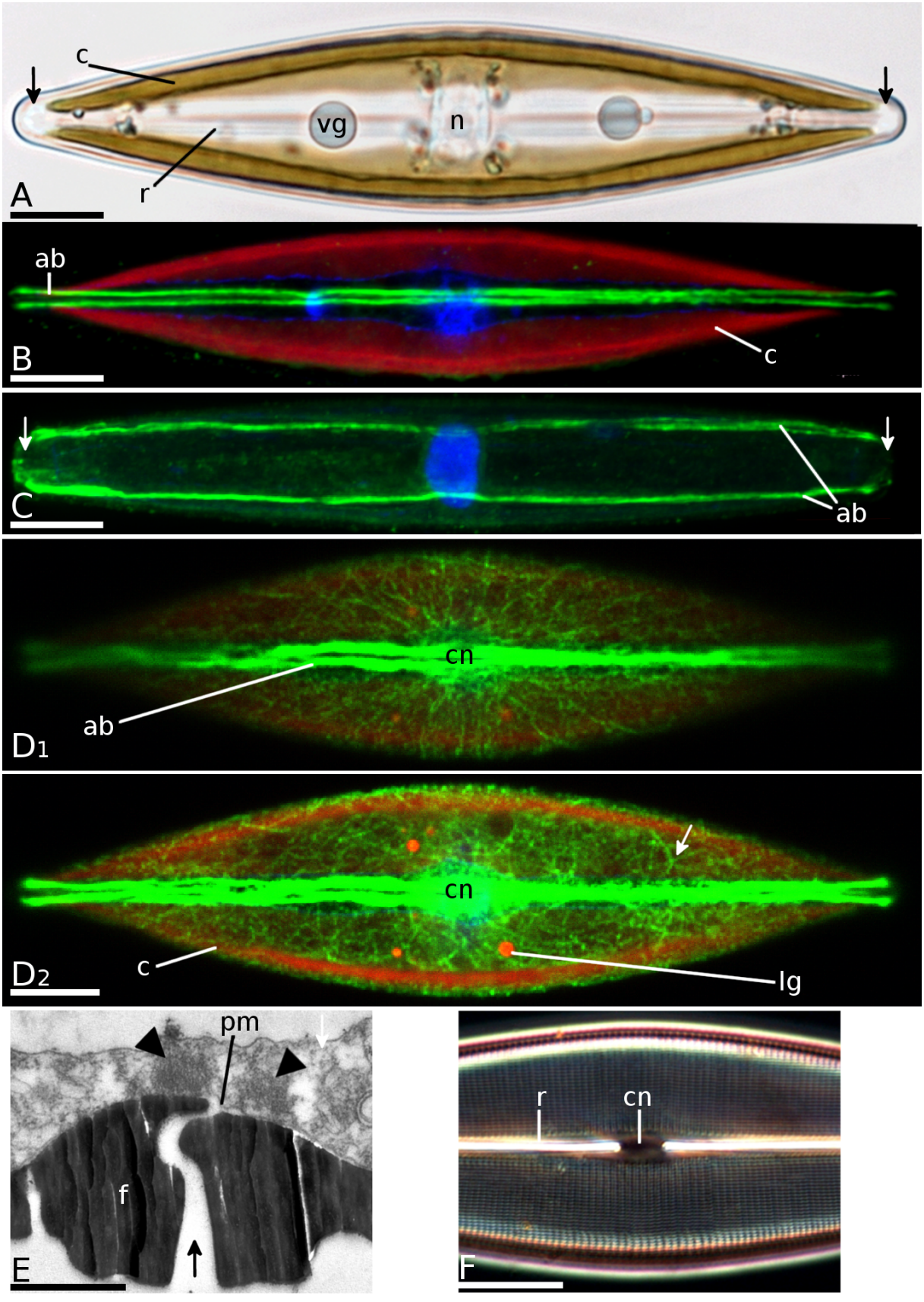
Cells treated for 6 h with the same concentration showed an even more severe effect, the fine cortical actin meshwork and the major cytoplasmic actin network fully disappeared (Fig. 7.2 B). Few filaments appeared in the cytoplasm predominately orientated along the longitudinal axis as well as few bundles near the girdle band region. Some of the bundles showed conspicuous branching. At the polar raphe endings, remnants of the two actin bundles merged with other actin fragments into a large aggregate (Fig. 7.2 B arrow). Another conspicuous aggregate formed at the the central nodule after a 6 h treatment. Within this region it was indistinguishable, whether the labeled structures were remnants of the thick actin bundles, or whether they originated from the actin aggregates.

After 24 h of jasplakinolide treatment the raphe actin bundles fully disappeared. A few actin filaments were still present throughout the cell, again longitudinal orientated and the actin aggregates at the central nodule and at the polar raphe ending were also still prominent although somewhat reduced (Fig. 7.2 C arrow).

7.3 Visualization of the MT cytoskeleton

The dense MT structure within *C. cuspidata* was very symmetrical, radially organized with the nucleus as a center of symmetry (Fig. 7.3 A, E), reflecting the high symmetry of the cell as shown in the bright field image (Fig. 7.1 A). MTs radiated out from the nuclear surface extending into the periphery of the cell. Although long MTs were spanning the cell, they only weakly interspersed the outermost cell poles (Fig. 7.3 B gray arrowhead, Appendix Fig. 18.1). This is in contrast to the actin cytoskeleton, which was strongly present in the cell poles. Double labeling of actin and tubulin revealed that the MTs near the raphe aligned left and right with the two raphe actin bundles (Fig. 7.3 A). The MTs appear to originate from several spots spread over the nuclear

Figure 7.1 (following page): The actin cytoskeleton in *C. cuspidata* shown by immunofluorescence labeling. A,B,D,F: Valve view of the cell; C: girdle band view. **A:** Bright field image of a *C. cuspidata* interphase cell with the two plastids (c) appearing as pale brown stripes in the girdle band region. The two prominent vacuoles on both sides of the centered nucleus (n) contain one volutin granules (vg) each. Near the cell poles (arrow) small vesicles are present. The raphe (r) along the longitudinal axis is slightly visible. **B:** Two parallel actin bundles (ab). **C:** Girdle band view. The two sets of actin bundles (ab) one on the hypovalve side and the other on the epivalve side are not connected (arrow). **D:** Two single optical sections of the same cell. D_1 captured just underneath the frustule showing the two thick parallel actin bundles (ab) and a fine actin meshwork radiating out form the area around the central nodule (cn). D_2 captured 1.6 μm further in the cell showing the actin network consisting of F-actin (arrow) and bundled F-actin. The highest actin network density is in the center of the cell under the central nodule (cn). **E:** TEM cross section of the raphe. The slit through the frustule (f) is called the raphe (arrow). The two actin bundles (black arrowheads) are close to the plasma membrane (pm). Scale bar 1 μm. **F:** Light microscopy image of a frustule, viewed under crossed pol-filters. The raphe (r) is clearly visible with the interrupting central nodule (cn). Blue — DAPI staining of the nucleus and volutin granule (v); green — actin; red — autofluorescence of the plastid (c) and lipid globules (lg). Scale bars 10 μm.


c
vg
n
r
A
ab
B
c
C
ab
cn
ab
D1
cn
c
lg
D2
pm
f
E
r
cn
F

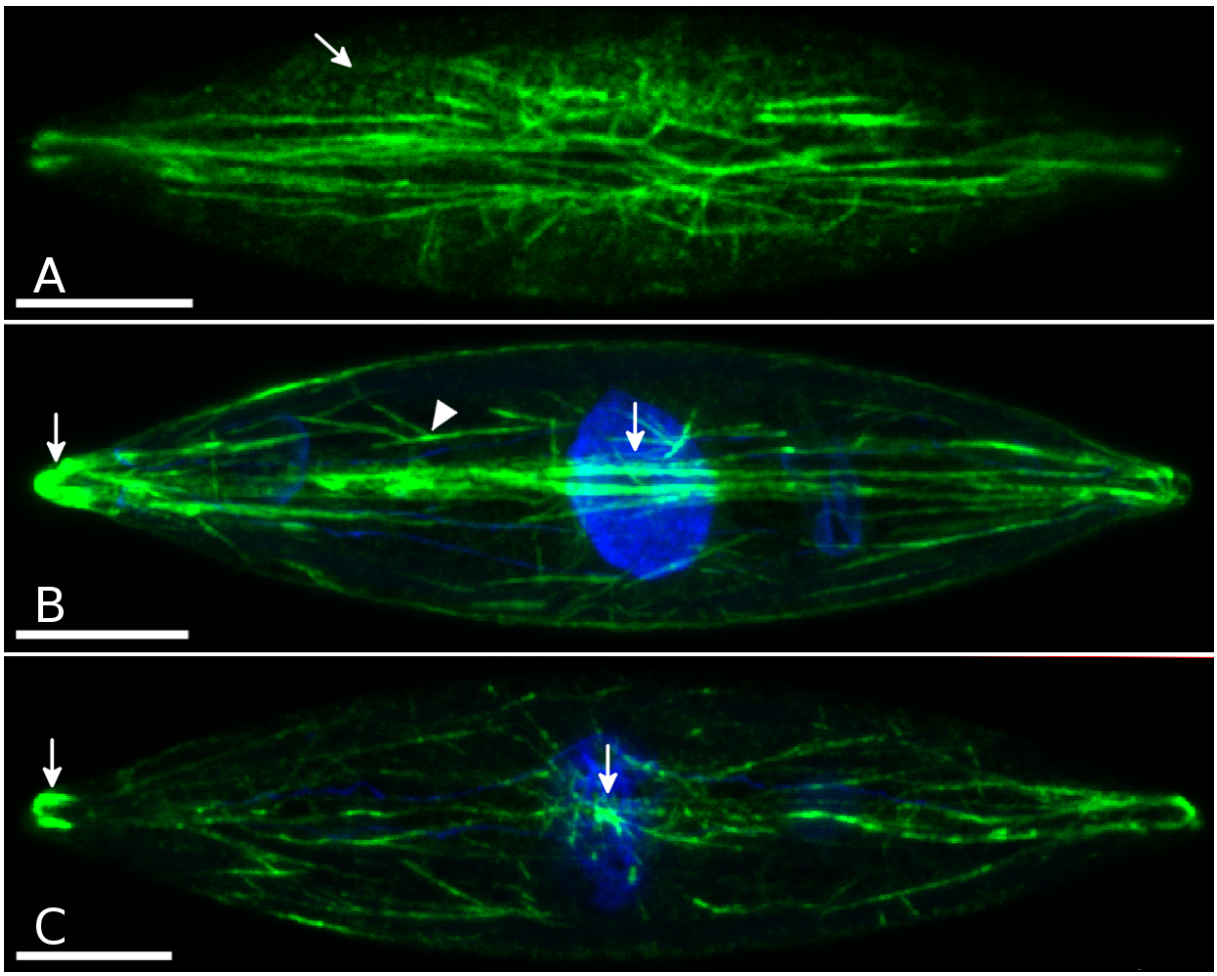


Figure 7.2: Jasplakinolide effects on the actin cytoskeleton. Cells treated with 1 μM jasplakinolide, shown in valve view. Projection of optical sections from the top of the cell about 4-5 μm inwards. **A:** Treatment with jasplakinolide for 1 h caused the two major raphe bundles to reorganize and partially disappear, but at the same time additional actin structures seem to be formed along the raphe (arrowhead). The fine cortical actin meshwork (arrow) is still present. **B:** After a jasplakinolide treatment of 6 h the fine cortical actin meshwork and the actin network fully disappeared and the actin bundles were only partly visible. Few straight and sparsely branched actin bundles are present (arrowhead). Actin aggregates (arrow) near the raphe ending and in the area of the central nodule. **C:** Cell treated for 24 h. Aggregate formation at the polar raphe ending (arrow), and in the area of the central nodule (arrow). green — actin; blue — nucleus. Scale bars 10 μm.

surface, often one of them is particularly prominent, at which several MTs converge in a radial fashion. This structure could be interpreted as the microtubule center (MC), (defined by Pickett-Heaps, 1991). The other smaller convergence points may be considered as additional MT organizing centers (MTOCs) and as such they should containγ-tubulin as an essential component (Brown and Lemmon, 2007). Indeed the γ-tubulin specific antibody marked more than a dozen spot-like structures on the nuclear surface (Fig. 7.4 A) and double labeling revealed that MTs originated from them (Fig. 7.4 B, arrowhead). In the center of the larger MC-like structure, a narrow bar was also labeled by the gamma-tubulin antibody (Fig. 7.4 B, arrow).

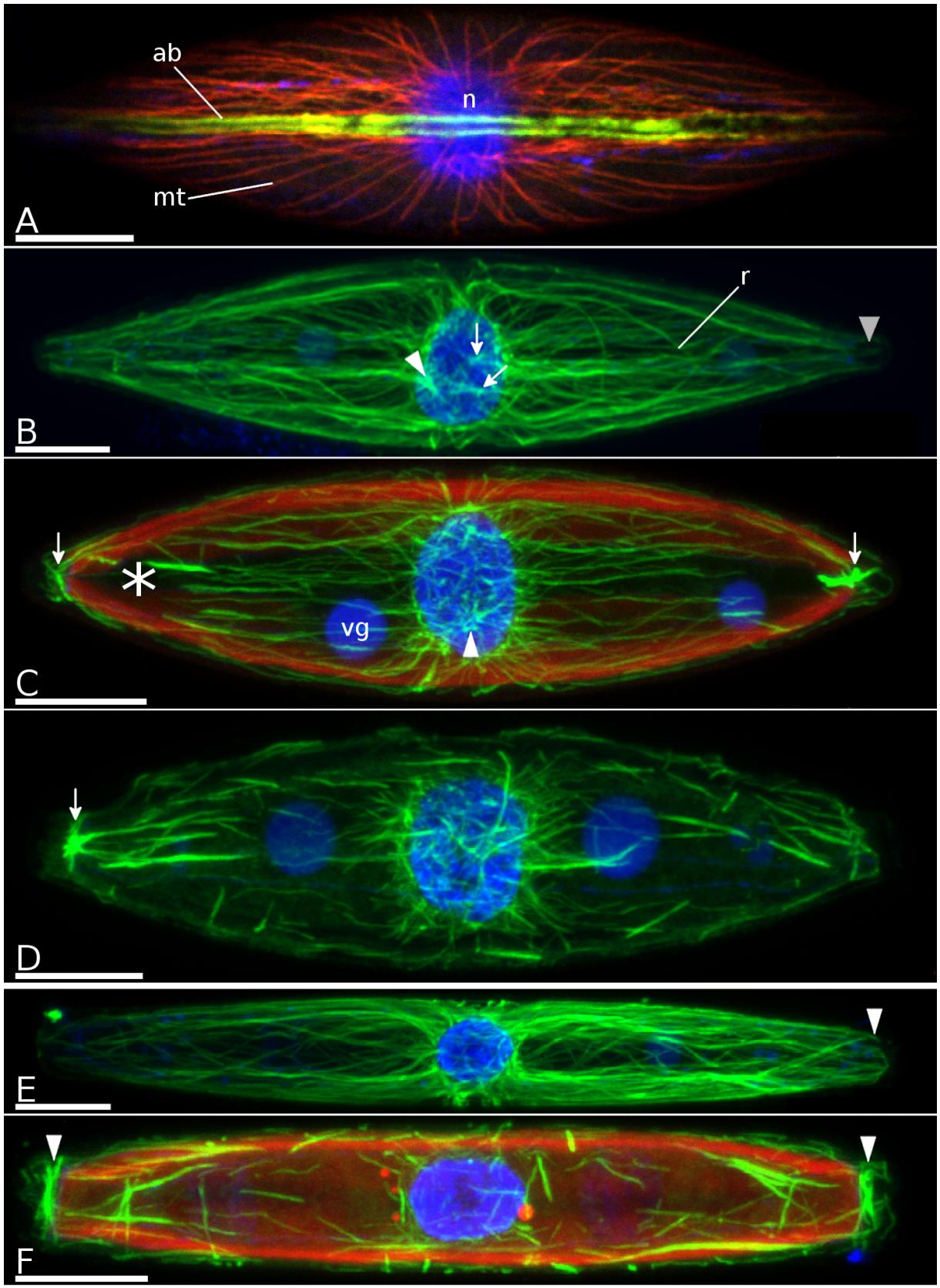
7.4 Effects of oryzalin on the MT cytoskeleton

A 0.5 μM oryzalin treatment for 1 h lead to depolymerization of the MTs. The long MTs shrank from the cell poles towards the nucleus (Fig. 7.3 C star), leading to an increased number of short MTs around the nucleus. The MC was still visible after oryzalin treatment (Fig. 7.3 C, arrowhead). Strongly labeled MT patches formed at the cell poles, just next to the plastid tips (Fig. 7.3 C, arrow).

After 1 h treatment with 1 μM oryzalin the MTs radiating symmetrically from the nucleus were depolymerized, and only short MTs around the nucleus remained, as well as MT fragments scattered around the cytoplasm. Few of the MTs near the cell poles appeared to radiate out from the aggregated MT patches (Fig. 7.3 D arrow). The girdle band view of cells reveals that the aggregated MT patches at the cell poles are vertically arranged just next to the tips of the plastids (Fig. 7.3 F arrowhead). Predominantly the MT patches appeared symmetrically at both cell poles, but occasionally patch formation at only one pole was observed (Fig. 7.3 D).

Longer treatments for 6 h with 1 μM of oryzalin reduced the MT patches into a cap-like formation of the plastid tips (Fig. 7.5 A, A'). These cap-like structures were located just around the area of the nucleoid. In contradiction to that, no MT structures were ever labeled at the plastid tips in untreated cells (Fig. 7.3 B, E). Furthermore γ-tubulin labeling of the cells did also not label the plastid tips (Fig. 7.4 A, arrowhead). Then again cap-like structures at the tips of the plastids were observed by TEM in untreated cells. This structures were tightly align to the plastid envelope, appeared filamentous and were not enclosed by a membrane (Fig. 7.5 B, C).

Figure 7.3 (following page): Effects of the MT drug oryzalin. A-D: valve view; E-F: girdle band view. **A:** Single optical section of a cell underneath the frustule. Double labeling of actin (green) and tubulin (red). B-F: Projection of optical sections of a cell. **B:** MTs radiate out from the MC (arrowhead), as well as from MTOCs (arrow). MTs spare the area underneath the raphe (r) and the cell poles (gray arrowhead). **C:** A 0.5 μM oryzalin treatment for 1 h depolymerized MTs. MTs near the cell poles (star) disappeared and a patch of accumulated MTs (arrow) at the tips of the plastids appeared. Arrowhead — MC. **D:** A 1 μM oryzalin treatment for 1 h made the MT patch structures even more prominent (arrow), some MTs were associated with this structure. **E:** Girdle band view of an untreated cell. Long MTs end in the poles of the cell (arrowhead). **F:** Girdle band view of a cell treated for 1 h with 1 μM oryzalin. Vertical MT patches (arrowheads) formed just next to the the tips of the plastids (red autofluorescence). Blue — nucleus (n) and volutin granules (vg); red autofluorescence — plastid and lipid granule. Scale bars 10 μm.


ab
n
mt
A
r
B
vg
C
D
E
F

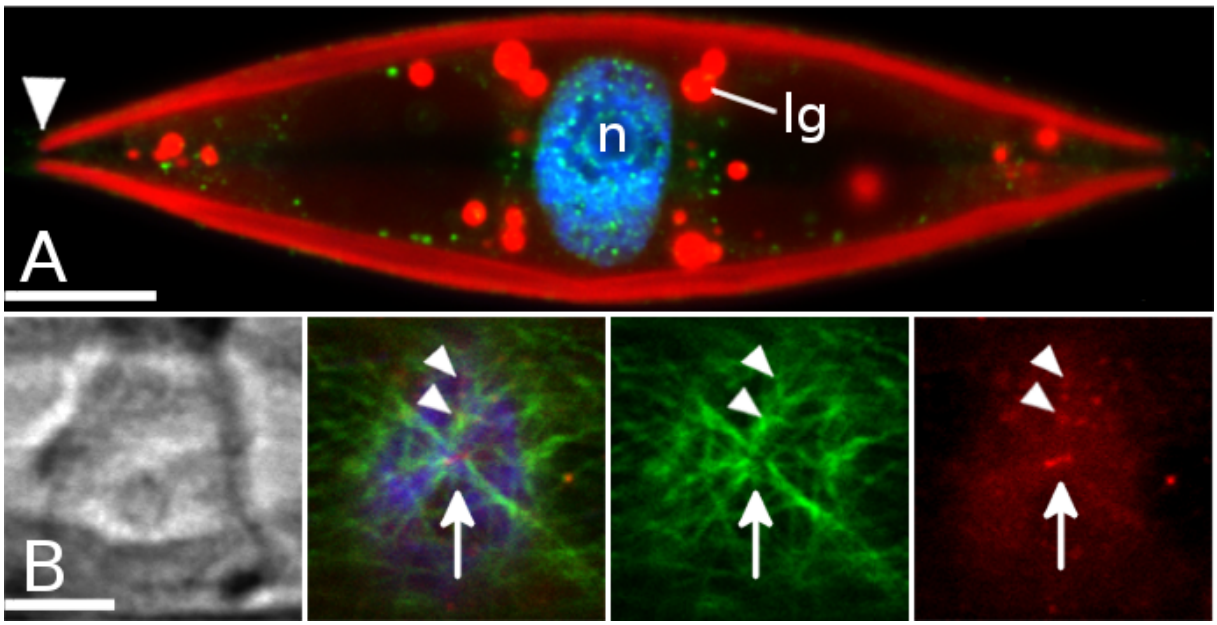


Figure 7.4: γ-Tubulin labeling. **A:** Projection of optical sections of the nucleus. γ-tubulin (green) is located around the nucleus in spot-like structures. No γ-tubulin labeling at the plastid tips (arrowhead). green — γ-tubulin; blue — DAPI staining; red — autofluorescence of the plastid (c) and lipid globular (lg); scale bar 10 μm. **B:** Optical sections of the nuclear surface. Double labeling of α-tubulin (green) and γ-tubulin (red), staining of DAPI in blue. MC (arrow) at the surface of the nucleus with γ-tubulin in the center. Spot-like structures resemble MTOCs (arrowhead). Scale bar 5 μm.

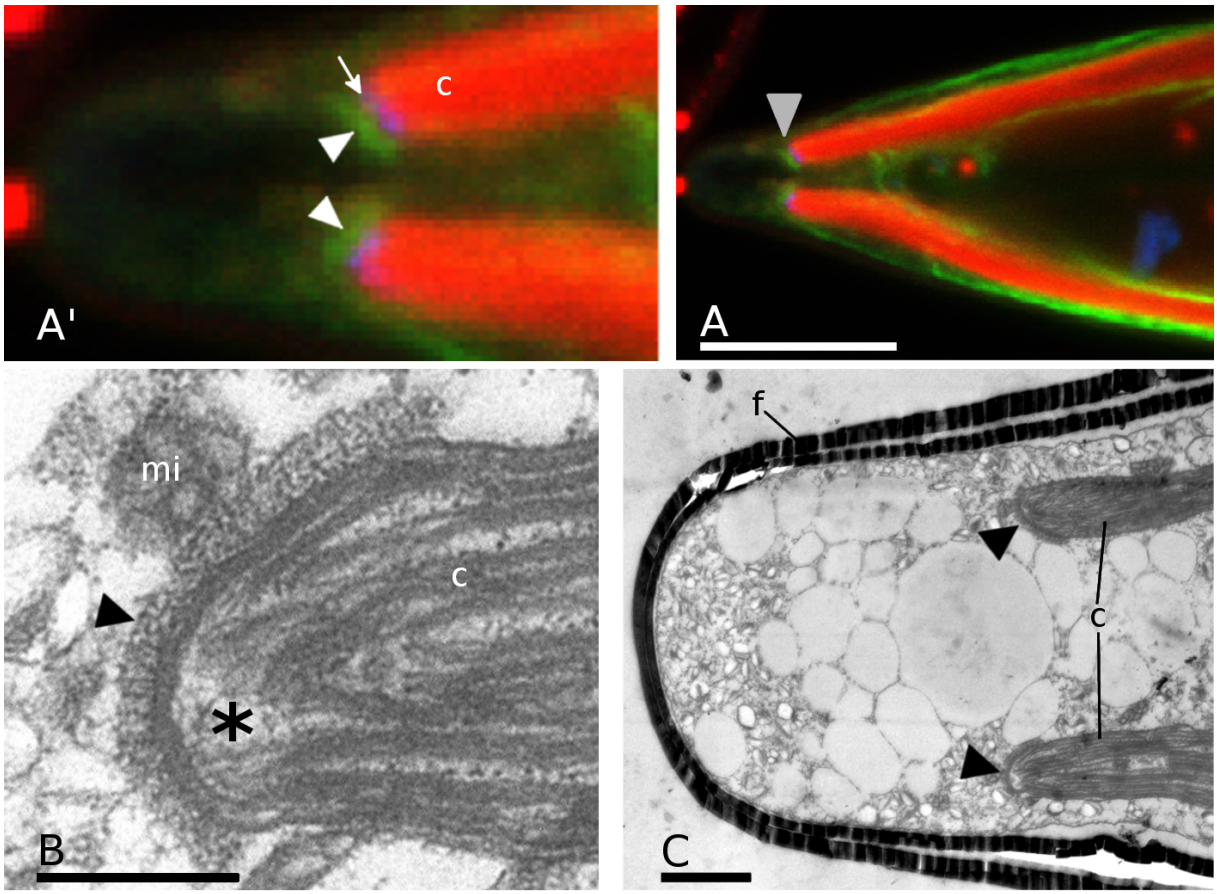


Figure 7.5: Cap-like structure at the plastid tips. **A:** A 1 μM oryzalin treatment for 6 h. Scale bar 10 μm **A':** Enlarged area of A under the gray arrowhead. The cap-like structure (arrowhead) is just next to the plastid tip where the nucleoid (blue) is located. Red — autofluorescence of the plastid; blue — nucleoid stained with DAPI; green — MT. **B:** TEM of an untreated cell. Cap-like structure (arrow) at the tip of a plastid (c). mi — mitochondria; *nucleoid; scale bare 400 nm. **C:** TEM overview of a longitudinal section of the cell pole with the two plastid (c) tips associated by a cap-like structure (arrowheads). f — frustule; Scale bare 400 nm.

8 Discussion

8.1 The actin cytoskeleton

In previous studies on the diatom cytoskeleton, actin filaments have been visualized using fluorochrome-conjugated phalloidin. In contrast to the properties of an antibody, this fluorescent probe only detects F-actin but not G-actin (Faulstich et al., 1988). Fluorochrome-conjugated phalloidin also has a tendency to reveal bundles rather than single filaments, such as the raphe actin bundles and the actin bundles at the leading edge of the SDV (Poulsen et al., 1999; Van de Meene and Pickett-Heaps, 2002; Tesson and Hildebrand, 2010a). Due to the exclusive use of fluorochrome-conjugated phalloidin, fine actin structures like the actin meshwork in *C. cuspidata* have not been observed in this detail before.

Actin inhibitor studies have investigated the effects of cytochalasins toward cytokinesis, valve morphogenesis, or the locomotion process depending on the raphe actin bundles (Poulsen et al., 1999; Pickett-Heaps and Spurck, 1982). Cytochalasins have been employed in countless studies in the past on the basis of their pharmacological interference with actin (Cooper, 1987). Yet in the diatom field, their efficiency as actin inhibitors is a quite controversial issue. Some authors claim that it also disrupts actin functions (Pickett-Heaps and Spurck, 1982; Van de Meene and Pickett-Heaps, 2004), whereas others did not observe any effects (Tesson and Hildebrand, 2010a) even at high concentrations of up to 100 μM and regardless of which of the available chemical variants (cytochalasins C, D, E, J, and H) had been used (Poulsen et al., 1999). As an alternative, Poulsen et al. (1999) used latrunculin A and reported it to be a potent actin depolymerizing drug at concentrations of 0.2-1 μM within 30 min for three different pennate diatoms. Latrunculin A could be confirmed in this study as a potent actin drug for *Cylindrotheca fusiformis* (chapter 11.2.2), however it did not show any affect towards *C. cuspidata*. Latrunculin A and also its stereoisomer latrunculin B did neither impair locomotion, nor did it depolymerize the actin cytoskeleton in *C. cuspidata* at concentrations of 1-5 μM up to 24 h. Therefore, jasplakinolide was tested for *C. cuspidata* as yet another alternative, and effectively disrupted the actin cytoskeleton already at a concentration of 1 μM for 1 h.

The efficiencies of actin drugs in diatoms appear to be quite different, varying from species to species even within the same class. It would be interesting to investigate, if this difference correlates with mutations in the inhibitor binding domains of the actin genes. Such an approach is now in reach because four diatom genomes have been sequenced and their actin genes are known (chapter 15).

Jasplakinolide promotes actin polymerization and stabilizes actin filaments, thereby influencing the critical concentration of the available G-actin poll. *In vivo* this leads to disrupted actin filaments and induces polymerization of monomeric actin into amorphous aggregates (Bubb, 2000; Lázaro-Diéguez et al., 2008; Holzinger, 2010), while stable actin structures are not as affected by

jasplakinolide as dynamic structures.

C. cuspidata cells treated with 1 μM jasplakinolide for 1 h still contained a cortical actin meshwork, whereas the major cytoplasmic actin structures became disrupted. Leading to the assumption, that the cortical actin meshwork in diatoms is not a very dynamic structure. All organisms undergoing cellular shape changes during growth, development, or in the context of locomotion, require a dynamic actin cytoskeleton in the cell cortex (reviewed in Blanchoin et al., 2014). In plants the abundance of the cortical actin array correlates with axial cell expansion. It is thought that actin arrays coordinate vesicle and organelle trafficking and thereby modulate the delivery of cell wall materials and cellulose synthesis machinery to the plasma membrane (Gutierrez et al., 2009; Crowell et al., 2009; Szymanski and Cosgrove, 2009; Sampathkumar et al., 2013). Due to their rigid, silicified cell wall diatoms are drastically restricted in cell shape changes. Cell movement involving the formation of lamellipodia and filopodia, like in animal cells, is just not possible. Furthermore gradual, polarized growth of the cell by maintaining a certain plasticity of the cell wall, like in plant cells, is also impossible. Also diatoms grow by axial expansion, the prefabricated frustule elements get inserted in-between the existing cell wall by a huge exocytosis process. Consequently a dynamic cortical actin is not as essential for diatoms as for other organisms. Even though diatoms do not form any variable cell protrusions, it is not precluded that they contain some proteins capable of connecting the actin cytoskeleton with the plasma membrane. Indeed, it is known from previous studies that both cytoskeletal elements, actin and MTs, are involved in the shaping of the SDV (Pickett-Heaps and Kowalski, 1981; Schmid, 1987; Pickett-Heaps et al., 1990; Schmid, 1994; Van de Meene and Pickett-Heaps, 2002, 2004; Kröger and Poulsen, 2008; Tesson and Hildebrand, 2010a,b). It is also very likely that the cortical actin in diatoms is connected to the plasma membrane. The cortical actin meshwork in *C. cuspidata* is predominantly radially organized arising from the area around the central nodule just above the nucleus. Golgi bodies are usually located within the perinuclear area in the diatoms (Werner, 1977), and this is also the case in *C. cuspidata* (data not shown). Secretory vesicles from the Golgi bodies destined for exocytosis have to be transported from the perinuclear area to the plasma membrane. Considering the stability of the cortical actin meshwork including its radial organization, it appears to be ideally suited to assume a distributing "railway function" for vesicular trafficking, from the locations of synthesis to the secretory target sites and the other way round, as well as between the sites of exocytosis and endocytosis.

The major cytoplasmic actin network in *C. cuspidata*, in comparison to the cortical actin seems to be more dynamic, as jasplakinolide induced a severe disruption in the organization within one hour. The structure of the raphe actin bundles partly and the major cytoplasmic actin network fully disappeared, while bundled filaments polymerized in the raphe region (6 h treatment). Treatments of 24 h seemed to disrupt the actin filament nucleation/polymerization along the raphe and clustered the actin aggregates at hotspots, always located at the polar raphe ending and in the region under the central nodule (Fig. 7.2). An explanation for the formation of longitudinally arranged bundled actin filaments could be, that jasplakinolide, based on secondary effects, reduces the efficiency of actin-bundling proteins within the thick actin bundles. Leading to a separation, splitting of the thick bundles into thinner bundles. A longer treatment would eventually depolymerize these filaments leading to the aggregate formation at the hotspots. However, nothing is known about potential effects of jasplakinolide on actin-bundling proteins. Another possibility would be, that new filaments nucleate along the longitudinal axis, underneath the raphe. This observation is consistent with reports that jasplakinolide increase the density of actin filaments adjacent to the

plasma membrane, while stress fibers are disrupted (Shurety et al., 1998; Patterson et al., 1999). Transferring this to the situation in *C. cuspidata* could mean, that the region underneath the raphe is an area of high actin nucleation activity. After a long treatment (24 h) these regions of nucleation were eventually disrupted and further formation of aggregation at hotspots located at the cell poles and under the central nodule was promoted. It has been previously shown for mammalian cells that the formation of large actin aggregations after jasplakinolide treatment is promoted by retrograde transport of smaller actin aggregates along the MTs towards the MTOC (Garcia-Mata et al., 2002; Lázaro-Diéguez et al., 2008). This is consistent with the formation of the actin aggregate under the central nodule in *C. cuspidata* possibly reflecting the MTOC, but the aggregates at the cell poles of *C. cuspidata* would remain unexplained.

The two thick raphe actin bundles, first described by Drum (1963), are an essential structure of cell locomotion in pennate diatoms. They are the best studied actin structures in diatoms, however, not much is known about the formation and dynamics of these actin bundles. Up to now it is unclear, if these bundles consist of actin filaments oriented in a parallel, anti-parallel, or mixed polarity fashion. Furthermore, it is still an unsolved question, if they nucleate at the cell poles and polymerize along the raphe, or if they nucleate from various points all along the raphe. Poulsen et al. (1999) previously demonstrated with a recovery experiment, that after latrunculin A depolymerization, the raphe actin bundles reappeared within a time span of under 10 sec, after the removal of latrunculin A. These bundles were spanning the cells with the length of about 50 μm, but were only slightly stained and appeared not as thick as in cells 10 min after drug removal. The fastest recorded actin elongation rates in dilute buffers are about 100–1000 subunits/s corresponding to about 0.3–3 μm/s (Pollard et al., 2000). Assuming that diatoms have an exceptionally fast actin elongation with 3 μm/s, would lead only to 30 μm within 10 sec and not to the reported 50 μm in under 10 sec (Poulsen et al., 1999).

Considering these calculations and the results that the area underneath the raphe has a high actin nucleation activity, it appears convincing that the raphe actin bundles nucleate all along the longitudinal axis underneath the raphe. It is likely that the raphe actin bundles after repolymerization consist of short bundled filaments, which might then further polymerize to longer filaments. This would be consistent with the observation of Poulsen et al. (1999), that the bundles were faintly stained at the moment of their reappearance and only after 10 min accumulated a sizable signal. Candidate proteins that would be able to mediate the nucleation of F-actin close to the membrane all along the raphe would be formins (Sagot et al., 2002a; Evangelista et al., 2002; Harris and Higgs, 2004; Chesarone et al., 2009) (for more detail see chapter 15.3.3). It would be therefore interesting, to experimentally investigate the localization of these formins in a future study and further investigate the possible interactions between formins and actin in diatoms.

8.2 The MT cytoskeleton

Microtubules have been described as components of the diatom cytoskeleton in fine structural research for more than 20 years (for a summary see Pickett-Heaps et al., 1990). The present study is the first account of the MT-cytoskeleton in a pennate diatom at mature interface stage. Considering the organization of the MT-system from the nuclear surface and its radial construction, the MT cytoskeleton of *C. cuspidata* is similar to reported MT structures of centric diatoms, but

the MT organization in *C. cuspidata* is much more symmetrical (Wordeman et al., 1986; Van de Meene and Pickett-Heaps, 2002, 2004; Pollock and Pickett-Heaps, 2005; Tesson and Hildebrand, 2010a,b). This study is also the first to label γ-tubulin in a diatom, and to co-localize MTs with γ-tubulin and describes the behavior of the MT-system in response to inhibitor treatment in a pennate diatom.

Examination of the perinuclear region demonstrates the presence of several smaller and one large center of MT convergence and these centers co-localized with γ-tubulin as would be expected for a genuine MTOC (Shu and Joshi, 1995; Brown and Lemmon, 2007). The larger center of MT convergence is likely to represent the MC (according to Pickett-Heaps, 1991; De Martino et al., 2009). The diatom-specific term MC refers to a globular structure close-by the nucleus, which is seen as a guiding structure for the intracellular placement of the nucleus. However, it has not been experimentally confirmed that this structure truly organizes MTs (Pickett-Heaps, 1991). Here the co-localization of the putative MC with γ-tubulin indicates, that this structure indeed is capable to organize MTs. Therefore the MC may be considered as a large MTOC with special additional features typical for the diatoms. In *C. cuspidata* MTs originate not only from the MC, but also from the small MTOCs co-localizing with the spot-like γ-tubulin structures at the nuclear surface. Multiple MTOCs at the nuclear envelope as shown here, have not been reported for diatoms before, but are described for higher plants (Schmit, 2002).

The effect of oryzalin on plant MTs has been originally described as a fast, concentration-dependent depolymerization, an order of magnitude more potent than colchicine, whereas it was almost ineffective on animal MTs (Morejohn et al., 1987). More recently, it has been discovered that oryzalin binds to the α-subunit of the tubulin heterodimer and thus inhibits polymerization. Gradually the pool of heterodimers containing oryzalin-bound α-tubulin increases due to normal turn over activity from polymer to heterodimer, until no more polymerization-competent heterodimers are left (Morrissette et al., 2004). Oryzalin has rarely been used as anti-MT agent in diatoms. A study on plasmolysis and recovery from plasmolysis was performed in the centric diatom *Ditylum* involving treatment with oryzalin (Pollock and Pickett-Heaps, 2005), but there has been no detailed study addressing the action of oryzalin on the complete interphase cytoskeleton in any diatom. It is shown here for the first time that this agent is a highly effective anti-MT drug in diatoms.

Cells treated with a concentration of 0.5 μM oryzalin have shortened MTs at the cell poles and a reduced MT network around the nucleus, but the MC seemed to be still present. At a concentration of 1 μM oryzalin the radially outgrowing MTs disappeared and only fragmented MTs, as well as a reduced MT network around the nucleus were present. It seems that a low concentration is shortening the MTs in *C. cuspidata* and that a higher concentration of 1 μM leads to further depolymerization, but that this concentration is not yet high enough to fully depolymerize all MTs. This observation is in agreement with previous reports for Apicomplexa, which showed that a low concentration of 0.5 μM oryzalin leads to shortened non-functional MTs, whereas only at an even higher oryzalin concentration of 2.5 μM all MTs had been depolymerized (Morrissette and Sibley, 2002).

In *C. cuspidata*, MT patches accumulated after oryzalin treatment at the tips of the plastids, an effect that has not been reported before. This result has been not described for oryzalin before and rather resembles the effect described for vinblastin, another MT inhibitor used in cancer therapy (Rai and Wolff, 1997). These MT accumulations at the plastid tips occurred only after oryzalin treatment and not in untreated cells, therefore it might be that this regions assume some sort of

a tethering/anchoring function for MT fragments. Another possibility is that the MT aggregation at the plastid tips is connected to structures relevant for plastid positioning, or that theses areas do contain further MTOCs within *C. cuspidata*. Pickett-Heaps (1991) previously reported that the MC is involved in plastid positioning preceding mitosis in diatoms, but for this event the MC is repositioned from the nucleus towards the middle of the plastid. The assumption that the plastid tips contain MTOC activity would also explain the observed originating/ending of MTs in the MT-patches. MTOCs located on the plastid surface haven been reported in early land plants, including γ-tubulin (Shimamura et al., 2004). However, since in *C. cuspidata* the γ-tubulin antibody did not label the area around the plastid tips, this specific region may not be considered as a genuine MTOC, although the ability for binding MTs or tubulin heterodimers, respectively, is evident. This interpretation is consistent with the finding that after prolonged incubation with 1 μM oryzalin for 6 h, the MT-patches shrink back to faintly labeled, though distinctive cap-like structures at the tips of the plastids. This finding prompted further electron microscopic examination of this cell region, revealing that these structures correlated perfectly well with fibro-granular cap-like structures tightly associated with the plastidal tips at the cell poles in untreated cells. Neither their composition, nor their function are currently known. It can only be speculated that these structures assume some sort of MT regulation, or anchoring of MT fragments within *C. cuspidata*, not connected to the function of γ-tubulin. Although, it cannot be excluded that there is a very special form of γ-tubulin in diatoms, not recognized by the antibody used here. One first step in a future study aimed at experimentally addressing the function of this structure would be an *in vitro* assay involving polymerization of MTs in the presence of isolated plastids or plastidal envelope membrane fractions, respectively. If this reveals an affinity between the two components, it should be worthwhile to be pursued further.

Transformation
in vivo visualization of actin

9 Introduction

The actin cytoskeleton is a highly dynamic network, reorganizing throughout the cell cycle and reacting towards extracellular stimuli. Immunofluorescence labeling of the actin cytoskeleton gives a very good, but static impression of the actin organization. In order to monitor real-time dynamics, a powerful methodology based on fluorescent protein fusion constructs has been established over the recent years. In many cases, a derivative of the green fluorescent protein (GFP) is translationally fused directly to the target protein (Snapp, 2005). But for actin it has been shown that such a fusion may not lead to the visualization of the actin cytoskeleton in the target cell, because the fused GFP impaired the function of actin (Westphal et al., 1997). Therefore, a more indirect approach has been taken, where GFP is not fused to actin but to an actin-binding protein or just to the its actin-binding domain. Several of those constructs based on a variety of actin-binding protein sequences have been developed (Higaki et al., 2007; Lemieux et al., 2014). One of these reporters is Lifeact, a 17 amino acids long peptide derived from yeast Abp140p (Riedl et al., 2008). Lifeact clearly labels filamentous (F-)actin as well as fine cortical structures in a wide range of organisms from fungi, over mammals to plants (Lichius et al., 2011; Era et al., 2009; Riedl et al., 2010). Additionally, Lemieux et al. (2014) claim that Lifeact is currently the best available probe to visualize the dynamic changes in actin networks. Lifeact fused to GFP has therefore been chosen in the present study as a fusion construct to visualize the diatom actin cytoskeleton *in vivo* for the first time.

Genetic manipulation methods for common model organisms, i.e., *Arabidopsis thaliana*, *Mus musculus*, or *Drosophila melanogaster*, are well established, however, for diatoms most of these methods are still in their infancy. Since the diatom research field is expanding, several groups successfully established genetic transformation protocols using the microprojectile bombardment method for different marine diatom species, including four pennate diatoms, *Phaeodactylum tricornutum* (Apt et al., 1996), *Fistulifera* sp. JPCC DA0580 (Muto et al., 2013), *Cylindrotheca fusiformis* (Fischer et al., 1999), and *Navicula saprophila* (Dunahay et al., 1995) as well as three centric diatoms, *Thalassiosira pseudonana* (Poulsen et al., 2006) *Cyclotella cryptica* (Dunahay et al., 1995) and *Chaetoceros sp.* (Miyagawa Yamaguchi et al., 2011). Until now, no fresh water diatom has been transformed. *P. tricornutum* is the most widely used model organism for diatom transformation and the only diatom, for which a protocol off genetic transformation by microprojectile bombardment and by electroporation is available (Zhang and Hu, 2013). Unfortunately the transformation efficiency by electroporation as well as by microprojectile bombardment is still rather low, though it is slowly getting improved (Miyahara et al., 2013).

Almost all of the established transformation protocols for diatoms use reporter constructs with an endogenous promoter and terminator to drive the expression of the gene of interest. Overexpressing promoters commonly used in other organisms, such as the viral 35S-CaMV plant promoter, do not drive gene expression within diatoms (Dunahay et al., 1995; Apt et al., 1996; Kroth, 2007). Few successful transformations are reported using a plasmid with a promoter derived from

another diatom genus, but within the same diatom lineage. Nevertheless, in diatoms, endogenous promoters result in a higher expression efficiency of the reporter gene, than exogenous promoters (Muto et al., 2013).

All diatom species that have been transformed so far, show a length from 5 μm to 10 μm. The lager diatoms, like *C. fusiformis* are elongate and thin with a diameter of about 5 μm. None of these organisms are ideal to study the organization, and dynamics of the actin cytoskeleton, because of their rather small size. Whereas *C. cuspidata* with a length of 100 μm and a width of 20 μm has been already proven by immunofluorescence studies (chapter 7) to be a suitable model organism to study the actin cytoskeleton. Therefore, it was the aim of the present study to transform *C. cuspidata* with a constructed LA-GFP expression cassette. Consequently, a transformation protocol by microprojectile bombardment had to be established for *C. cuspidata*, including an expression cassette containing an endogenous promoter/terminator. However, since the transformation of *C. cuspidata* did not succeed within this study, *C. fusiformis* was transformed with a constructed LA-GFP expression cassette driven under an endogenous promoter/terminator. Although *C. fusiformis* is rather small to study the dynamics of the actin cytoskeleton, it is still possible to observe the actin architecture and its dynamic. was also transformed with a constructed LA-GFP expression cassette.

Table 10.1: Vectors used in this study. FCP — fucoxanthin chlorophyll binding protein; p — promoter; t — terminator.

Name	Sequences used	Reference
pBluescript II Sk - (BSK-)		Prof. Menzel, University of Bonn
493	eGFP	N. Poulsen, B Cube, University of Dresden
333	FCPp-FCPt *C. fusiformis*	N. Poulsen, B Cube, University of Dresden
334	FCPp-Ble-FCPt *C. fusiformis*	N. Poulsen, B Cube, University of Dresden
353	FCPp-NatII-FCPt *C. fusiformis*	N. Poulsen, B Cube, University of Dresden
370	TP-FCPp-NatII-FCPt *T. pseudonana*	N. Poulsen, B Cube, University of Dresden
P2-2	FCPpb-NatII-FCPt *P. tricornutum*	P. Kroth, University of Konstanz
pPHA-T1-nptII	FCPp-neo-FCPt *P. tricornutum*	P. Kroth, University of Konstanz
StuI-GFP-pPhaT1	eGFP, ble *P. tricornutum*	P. Kroth, University of Konstanz

10 Material & Methods

10.1 Solutions & Enzymes

All molecular work and preparation of media was done according to the protocols of Sambrook et al. (2001). Enzymes were used accordingly to the manufacturers advices (Tab. 17.1).

10.2 DNA

10.2.1 Vectors

The following table contains the plasmids used for transformation and the underlying vectors, which were used to generate all constructs in this study. Desired sequences were either excised by restriction digestion, or amplified by PCR reactions.

10.2.2 Primer

Table 10.2 shows the primers used in this study. If primers were used to amplify fragments designated for molecular cloning, respective restriction sites were added.

Table 10.2: Primers used in this study. Small letters indicate added restriction sites.

Name	Purpose	Sequence (5'→3')
To amplify an actin gene of the *C. cuspidata* genome		
AF1	Degenerated Actin (for) 1	TGGCACCACACYTTCTACAA
AF2	Degenerated Actin (for) 2	ATGGARAAGATCTGGCACCACAC
AR1	Degenerated Actin (rev) 1	CCVGACTCATCGTACTC
AR2	Degenerated Actin (rev) 2	GGGWACATGGTGGTWCC
AR3	Degenerated Actin (rev) 3	GGCTGGAAVAGVACCTC
Inverse PCR to amplify the promoter/terminator of the actin genes		
ActAIF1	Inverse PCR ActinA (for)	AGTGTGGGTTACACCATCACC
ActAIF2	Inverse PCR ActinA (for)	GGTGATGGTGTAACCCACACT
ActAIR1	Inverse PCR ActinA (rev)	GTTACTCTTTGACCACCACGG
ActAIR2	Inverse PCR ActinA (rev)	CCGTGGTGGTCAAAGAGTAAC
ActBIF	Inverse PCR ActinB (for)	GTGAGTAACACCATCACCCGA
ActBIF2	Inverse PCR ActinA (for)	TCGGGTGATGGTGTTACTCAC
ActBIR1	Inverse PCR ActinB (rev)	CCGCTGGTATTGACCAAACCA
ActBIR2	Inverse PCR ActinB (rev)	TGGTTTGGTCAATACCAGCGG
Positive control of the Inverse PCR method		
SIF	Stonin *A. thaliana* (for)	GCTTCCATGCTGGTTGCTTAAGTTC
SIR	Stonin *A. thaliana* (rev)	CATACCTGTACAAAGGAAAG
Amplify a ubiquitin gene of the *C. cuspidata* genome		
UF1	Degenerated Ubiquitin (for)	AGCARCTBGARGATGGMCGWAC
UF2	Degenerated Ubiquitin (for)	ACRGGWAAGACCATCAC
UF3	Degenerated Ubiquitin (for)	AAGACCATCACBYTHGATGTCG
UF4	Degenerated Ubiquitin (for)	AAGATYCARGAYAAGGARGG
UF5	Degenerated Ubiquitin (for)	TKGARGATGGMCGTACMCT
UF6	Degenerated Ubiquitin (for)	AARATHCARGAYAARGARGG
UF7	Degenerated S27a (for)	ATGCARATYTTYGTVAARACYYTNAC
UF8	Degenerated S27a (for)	ACYYTNACRGGWAAGACCATCAC
UF9	Degenerated S27a (for)	AAGACCATCACBYTHGATGTCG
UR1	Degenerated S27a (rev)	GCCRTVAAGCABMMRAYBCC
UR2	Degenerated S27a (rev)	GCCATGAAGCAKCCAATBCC
UR3	Degenerated S27a (rev)	CSDYTTYTTYTTNCGYTTYTT
UR4	Degenerated S27a (rev)	CKCTTSRCCTTCTTKGG
UR5	Degenerated S27a (rev)	TTTCCRCAGTAGTAMCHATC
UR6	Degenerated S27a (rev)	GAAGRGTRGAYTCYTTYTGRATGTTG

UR7	Degenerated S27a (rev)	TTCTTDCCHCCWCCWCGHAGACG
UR8	Degenerated S27a (rev)	TTYTTVCKYTTCTTDCCHCCWCC

Amplify the native ubiquitin promoter/terminator of the *C. cuspidata* genome

Adaptor	GTAATACGACTCACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGT	
	Adaptor short 3'–5'	3′ H_2N-CCCGACCA-PO_4 5′
	Adaptor short 5'–3'	5′ PO_4-ACCAGCCC-N_2H 3′
Ap1	Adaptor Primer 1	GTAATACGACTCACTATAGGGC
Ap2	Nested Adaptor Primer 2	ACTATAGGGCACGCGTGGT
Ub1T1.1	Ubiquitin 1 Terminator1	ATGCAGATCTTCGTGAAAACC
Ub1T1.2	Ubiquitin 1 Terminator1 nested	CTTCGTGAAAACCCTGACAGG
Ub1T2.1	Ubiquitin 1 Terminator2	GGAAAGACGATCACGCTCG
Ub1T2.2	Ubiquitin 1 Terminator2 nested	GATCACGCTCGATGTTGAGC
Ub1P1.1	Ubiquitin 1 Promoter1	TTCTTTCTGGATGTTGTAGTCC
Ub1P1.2	Ubiquitin 1 Promoter1 nested	GTTGTAGTCCGAGAGAGTGC
Ub1P1.3	Ubiquitin 1 Promoter1 nested	TGCGTCCATCTTCCAACTGC
Ub2T1.1	Ubiquitin 2 Terminator1	ATGCAGATCTTCGTCAAGACC
Ub2T1.2	Ubiquitin 2 Terminator1 nested	CTTGACTGGAAAGACAATTACC
Ub2T1.3	Ubiquitin 2 Terminator1 nested	TGTCGAGCCTTCTGACACC
Ub2P1.1	Ubiquitin 2 Promoter1	ACCAAGTGAAGAGTAGACTCC
Ub2P1.2	Ubiquitin 2 Promoter1 nested	ACTCCTTTTGAATGTTGTAGTCC
Ub2P1.3	Ubiquitin 2 Promoter1 nested	AGTCCGACAAAGTGCGTCC

Construction of a expression cassette containing the native ubiquitin promoter/ terminator of *C. cuspidata*

PF	Ubi. promoter (for)	CTATAGGGCACGCGTGGTCGACG
PlR	Ubi. promoter long (rev) - StuI	aggcctGGCGGCATCTATTAGAATTG
PsR	Ubi. promoter short (rev) - StuI	aggcctAGGCCTGGCGGCCAAACATGC
TF	Ubi. terminator - 5' TAA (for)	AATTAAGTGATTGTAAGCCGCTG
TR	Ubi. terminator (rev)	TCCTTAGCAAAGTGTGAAGATG
pBlPT mutation - EcoRV, methylation (for)		CTGATATCGGCGGCTATCTATTAGAATTG
pBlPT/pBsPT mutation - EcoRV, methylation (rev)		CGCCGATATCAGGCCTATCGAATTCCT GCAGC
pBsPT mutation - EcoRV, methylation (for)		CTGATATCGGCGGCCAAACATGCC
NatF	natII (for)	ATGACCACTCTTGACGACACG
NatR	natII (rev)	TCAGGGGCAGGGCATG
NPTF	neo (for)	CGCTGCAGATGATTGAACAAGATG
NPTR	neo (rev)	CGCTGCAGTCAGAAGAACTCGTC
GFPF	eGFP (for)	attgatatcATGGTGAGCAAGGGCGAG
GFPR	eGFP (rev)	TTACTTGTACAGCTCGTCCATGC

Construction of the LA-GFP expression cassette to transform *C. fusiformis*

LAF	Lifeact with EcoRV, KpnI (for)	gatcgatatcATGGGCGTCGCTGACCTCAT

		CAAGAAGTTCGAGTCCATCTCCAAGGAGG AGggtaccactac
LAR	Lifeact with EcoRV, KpnI (rev)	ctagtggtaccCTCCTCCTTGGAGATGGA CTCGAACTTCTTGATGAGGTCAGCGACGCC CATgatatcgatc
GFPEcoRVF	eGFP - EcoRV (for)	attgatatcATGGTGAGCAAG
GFPNotIR	eGFP - NotI (rev)	attgcggccgcTTACTTGTACAG
PLGTF	Promoter LA eGFP Terminator - BamHI(for)	attggatccCTTTGTATTCAAAATAAG
PLGTR	Promoter LA eGFP Terminator - BamHI (rev)	attggatccTGGAAGCATGAAGTAC

Sequencing primer

M13F	BSK- sequencing primer M13 (for)	CAGGAAACAGCTATGACC
M13R	BSK- sequencing primer M13 (rev)	TGTAAAACGACGGCCAGT
SLAF	Lifeact screening (for)	ACCTCATCAAGAAGTTCG
SLAR	Lifeact screening (rev)	TCTCGTTGGGGTCTTTGC

10.2.3 Software

DNA *in silico* work to design constructs was done, using the DNA editing software GENtle for Linux (1.9.3), developed and maintained by Magnus Manske at the University of Cologne, Germany.

10.3 Molecular biological methods

All vectors were constructed using standard cloning techniques. Sequences of interest were PCR-amplified using either Phusion®, iProof or Omni polymerase with 3' → 5' exonuclease activity. Amplified fragments were purified by the Hi-Yield Gel/PCR DNA Fragments Extraction – Kit.

Fragments and vectors were cut by type II restriction endonucleases. Agarose gel electrophoresis was done to separate fragments and PCR-products of different sizes. Desired fragments were excised from the gel and purified by the Hi-Yield Gel/PCR DNA Fragments Extraction – Kit. Prior to ligation, fragments were purified using the Hi-Yield Gel/PCR DNA Fragments Extraction – Kit. Blunt-end cloning was preferably used (Tab. 10.3). In the cases of sticky-end DNA ligation performance, the digestion was incubated at either 16 °C over night, or at room temperature (RT) for 3 h, using T4 ligase. NEB Turbo Competent *E. coli* cells were transformed with the resulting plasmids by heat shock transformation.

Transformed cells were selected on LB agar containing 20 mg/ml ampicillin and when required with additional blue and white screening. Grown colonies were first analyzed by colony PCR. Positive colonies were cultured in liquid LB medium over night, followed by plasmid isolation using the NucleoSpin® Plasmid kit. The resulting constructs were verified by restriction digestion and if correct, sent for sequencing (Source Bioscience, Berlin, Germany). DNA was quantified by spectrophotometric measurement.

Table 10.3: Combined digestion and blunt-end ligation protocol. The restriction enzyme buffer was used and in order to optimize the condition for the T4 ligase reaction, ATP/DDT was added. MQ — ultrapure water from MilliQ®

Components	Amount in μl	
Vector	100 ng	
Insert	3 x (Vector ng x Insert bp)/Vector bp	
Buffer (enzyme)	2	
Enzyme	1	
T4	1	
PEG 4000	2	
ATP/DTT	2	Temperature -Time
MQ add to	20	at 21 ℃ over night

10.4 Diatom work

10.4.1 Preparation of diatoms for microprojectile bombardment

The cells were harvested for transformation when growth was at exponential phase. *C. cuspidata* cells were separated from the mucilage two days prior to transformation by rotation (described in chapter 2.4). Immediately before transformation, cells were concentrated by centrifugation: *C. fusiformis* and *P. tricornutum* 3,000 g for 10 min; *C. cuspidata* 3,000 g for 2 min. The final applied cell counts were 5x10^7/plate for *C. fusiformis*, 1x10^8/plate for *P. tricornutum*, and 1x10^7/plate for *C. cuspidata*. The cells were resuspended in 250-300 μl liquid culture medium per agar plate and spread in the middle of the agar plate in a 5.2 cm diameter circle (Fig. 10.1). After spreading, the agar plate was left open under the sterile bench for 10 min allowing the liquid culture medium to evaporate. The transformation was carried out after a combined protocol of the laboratory of N. Kröger and P. Kroth (personal communication).

Figure 10.1: Mounting of target cells. Diatoms were spread on the agar plate in the radius of the circle.

10.4.2 Transformation by microprojectile bombardment

The transformation was carried out under sterile conditions. All components were cleaned thoroughly with Ethanol and air dried in the clean bench prior usage.

The microcarrier was placed into the microcarrier holder and a 10 µl drop of well mixed coated-tungsten suspension was pipetted onto the center of the microcarrier and allowed to air dry. The rupture disk was placed into the rupture disk holder. The assembled microcarrier holder unit including the stopping screen was placed into the top slot of the biolistic system (PDS-1000/He, BioRad). The agar plate carrying the diatoms was positioned in the slot below. The air pressure inside the chamber was 28 psi. Diatoms were bombarded at rupture pressures of 1100 psi, 1350 psi and 1550 psi and shooting was repeated four times for each plasmid.

After transformation the cells of *P. tricornutum* and *C. fusiformis* were incubated over night at continuous light with 75 μE $m^{-2}s^{-1}$ for recovery. *C. cuspidata* was kept at 45 μE $m^{-2}s^{-1}$. The next day, 700 μl culture medium per plate was added to the cells and the cells were scraped of the plates. The sample from one shot was distributed and spread well onto four plates containing the antibiotics for selection of transformants (Tab. 10.5). After transformation *C. fusiformis* and *P. tricornutum*

Table 10.4: Coating of tungsten-particles with plasmid DNA.

Preparing the tungsten-particles	
Re-suspend	60 mg tungsten-particles in 1 ml Ethanol
Vortex	3-5 min
Centrifuge	1 min, high speed
Wash	pellet 3 times with 1 ml ddH_2O
Re-suspend	pellet in 1 ml ddH_2O
Aliquots	divide in 50 μl aliquots
Store	-20 ℃
Coating of tungsten-particles	
Under constant vortexing add to each 50 μl aliquot:	
Plasmid DNA (1 µg/µl)	5 µl
2.5 M $CaCl_2$	50 µl
0.1 M Spermidine	20 µl
Vortex	3 min, continuously
Centrifuge	10 sec, remove supernatant
Re-suspend	pellet in 250 µl Ethanol
Vortex	1 min
Centrifuge	10 sec, remove supernatant
Re-suspend	pellet in 50 µl Ethanol
Store	on ice until use

were grown on selection plates at continuous light (75 μE $m^{-2}s^{-1}$), whereas *C. cuspidata* was kept in a long day/night cycle (45 μE $m^{-2}s^{-1}$).

Table 10.5: Antibiotics used in this study to select transformants after microprojectile bombardment. Corresponding resistance genes are listed.

Antibiotic	**Species**	**Concentration**	**Resistance gene**
Nourseothricin (ClonNat)	*C. cuspidata*	2 mg/l	natII
Geneticin (G418)	*C. cuspidata*	25 mg/l	neo
Zeocin	*P. tricornutum*	75 mg/l	ble
Zeocin	*C. fusiformis*	1000 mg/l	ble

10.5 Transformation of *C. cuspidata*

10.5.1 Transformation of *C. cuspidata* with established diatom transformation plasmids

Before transforming *C. cuspidata* by microprojectile bombardment, it had to be tested if the cells survive the low pressure exposure. Further more it was assured that the pressure achieved with the rupture disks was strong enough to accelerate the tungsten particles to such an extend, that they penetrated the frustule.

In a first approach *C. cuspidata* was transformed with plasmids, that had already been successfully used to transform other diatoms. All of these expression cassettes are driven under endogenous

promoters of fucoxanthin chlorophyll binding proteins (FCPs). The plasmid 370 with a ClonNat resistance (natII) under an endogenous FCP promoter/terminator of *T. pseudonana*, the plasmid 353 with natII under an endogenous FCP promoter/terminator of *C. fusiformis*, the plasmid P2-2 with natII and the plasmid pPHA-T1-nptII with a G418 resistance (neo), both under an endogenous FCP promoter/terminator of *P. tricornutum* were used to transform *C. cuspidata* (Tab. 10.1).

Every plasmid was used in at least four different experiments on different days. Within every experiment each plasmid was shot four times, once at 1100 psi and 1550 psi and twice at 1350 psi, leading to a total of at least 16 shoots per vector at different pressures.

To optimize the recovery condition after the transformation procedure for *C. cuspidata*, different culturing conditions and recovery times were tested. Cells were recovered on agar plates, on agar plates with 2 ml medium added, and in liquid medium all for one to three days, before transferring them onto the selection plates containing the antibiotics.

As a positive control *P. tricornutum* was transformed with the StuI-GFP-pPhaT1 plasmid, to ensure that the transformation method was performed correctly.

10.5.2 Transformation of *C. cuspidata* with an endogenous promoter/terminator

Most diatoms so far have been transformed with a selectable marker/reporter gene construct expressed under an endogenous promoter. The present study therefore attempted to construct an expression cassette containing a new endogenous promoter and terminator.

Genomic DNA extraction of *C. cuspidata*

Genomic DNA of *C. cuspidata* was extracted according to the CTAB extraction protocol (modified after Murray and Thompson, 1980). Prior genomic DNA extraction, diatoms have been washed as described in chapter 2.4. Mortar and pestle were cooled down to liquid nitrogen temperature before usage.

- 700 µl pellet of centrifuged *C. cuspidata* cells were rapidly cooled down with liquid nitrogen and pestled.
- Fine diatom powder was dispersed in 5 ml extraction buffer (60 °C) and incubated for 15 min at 65 °C.
- 1 Vol Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v) was added, followed by incubation for 15 min on a slow rotating shaker. The sample was carefully invert every 5 min.
- Centrifuged (1600 x g) for 10 min at 20 °C.
- The aqueous phase was transferred into a clean tube and 1 Vol Phenol:Chloroform:Isoamyl Alcohol was added. The tube was carefully inverted 3 times.
- Centrifugation (1600 x g) for 10 min at 20 °C.
- Aqueous phase was transferred into a clean tube and 0,8 Vol Isopropanol was added. The tube was carefully inverted and incubated for 10 min at RT.
- Centrifugation (1600 x g) 15 min at RT.

- Supernatant was discarded and pellet washed with 50 μl 80% EtOH
- Centrifugation (1600 x g) 5 min at RT.
- Washing and centrifugation was repeated.
- EtOH was discarded and pellet air dried.
- Dry pellet was resuspended with 10 mM Tris-HCL EDTA pH8.
- Sample was RNase digested and genomic DNA was stored at 4 ℃.

Amplifying a ubiquitin gene

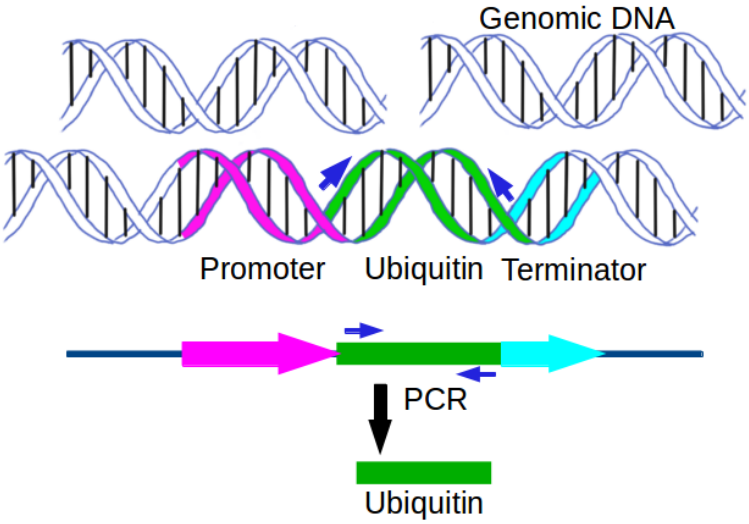


Figure 10.2: Amplifying an unknown sequence. To PCR-amplify the gene of interest of the unsequenced *C. cuspidata* genome, degenerated primers were designed to a well conserved gene. An example is shown here for a ubiquitin gene.

The conserved Ubiquitin-40S ribosomal S27a fusion gene was chosen to amplify a ubiquitin sequence of the unsequenced *C. cuspidata* genome (Fig. 10.2. The degenerated forward primers were designed within the 5' ubiquitin region and the degenerated revers primer within the 3' S27a region. All nine degenerated forward primers were tested with all eight degenerated revers primers to amplify a ubiquitin sequence (PCR condition Tab. 10.6).

Table 10.6: PCR program and component concentrations used to amplify a ubiquitin gene out of the unsequenced C. cuspidata genome with degenerated primers.

Component	Amount in μl	**Cycle - parameters**	Time
DNA	0.025	1) 98 ℃	2 min
Buffer (GC)	2	2) 98 ℃	10 sec
Phusion®	0.1	3) 56 ℃	20 sec
dNTP	0.2	4) 72 ℃	20 sec
Prim For	0.75	5) 72 ℃	10 min
Prim Rev	0.75	Repeat cycle 2-4 for 34 times	
MQ	5.2		

Table 10.7: Blunt-end digestion - Adaptor ligation for genome walking. *C. cuspidata* genomic DNA was digested with EcoRV, PvuII, StuI and the double-stranded adaptor sequences were simultaneously blunt-end ligated.

Components	**Amount** in μl
DNA	2,5 μg
double-strand adaptor	25 µM
Buffer (enzyme)	1,5
Enzyme	1,5 (10 units/μl)
T4	0,5 (6 units/μl)
Buffer (T4)	1,5
ATP/DTT	1
MQ add to 20 μl	-
	at 20 ℃ for 16 h

The PCR products of the primer pairs UF7-UR6 and UF7-UR7 were excised from the gel and purified. The purified PCR products were blunt-end ligated into the BSK- vector. *E. coli* was transformed with the two vectors and the colonies were selected by blue and white screening. 24 white colonies of each transformed line were further analyzed by Colony-PCR. The vectors of positive clones were isolated and sent for sequencing.

Genome walking - Amplifying an endogenous promoter/terminator of *C. cuspidata*

To amplify the endogenous ubiquitin promoter and terminator sequence, the genome walking method was chosen (Siebert and Chenchik, 1995; Siebert et al., 1995). The genomic DNA was digested with three different restriction enzymes, each recognizing a 6-base site, leaving blunt-ends (EcoRV, PvuII and StuI) (Tab. 10.7). An adaptor was designed to be ligated to the blunt-ends of the digested genomic DNA. The adaptor and the adaptor primers design closely followed the Genome Walker^{TM} Human Kit (Clontech Laboratories, Inc). The single-stranded adaptors were annealed at 95 ℃ for 2 min followed by a slow cooling to 25 ℃ over 60 min. Double-strand adaptors were stored at -20 ℃ until they were ligated to the digested DNA (Tab. 10.7).

The digested, adaptor-ligated genomic DNA fragment (DL-DNA) was the template to amplify the promoter/terminator sequence of the ubiquitin gene (Fig. 10.3). For a negative experimental control genomic DNA, as well as genomic DNA processed exactly like the DL-DNA just without the restriction enzymes as components (L-DNA) were used as a template. To amplify the endogenous promoter/terminator sequence of ubiquitin two consecutively PCR reactions were carried out (Tab. 10.8).

The PCR products of interests were excised from the gel and purified. The purified PCR products were cloned into the BSK- vector (Fig. 10.4). Thereafter *E. coli* was transformed with these vectors (pBWAlkProm, pBWAlkTerm) and colonies were screened by blue and white screening. White colonies of each transformed *E. coli* line were further analyzed by colony PCR. The vectors of positive clones were isolated and sent for sequencing.

Construction of a plasmid with an endogenous promoter/terminator cassette

A 489 bp long ubiquitin terminator sequence was PCR amplified from the generated pBWAlkTerm vector (Fig. 10.4) and subcloned into a BSK- vector at the SmaI cutting site resulting in the pBT vector (Fig. 10.5). A ubiquitin promoter sequence of 583 bp length, as well as a shorter variance

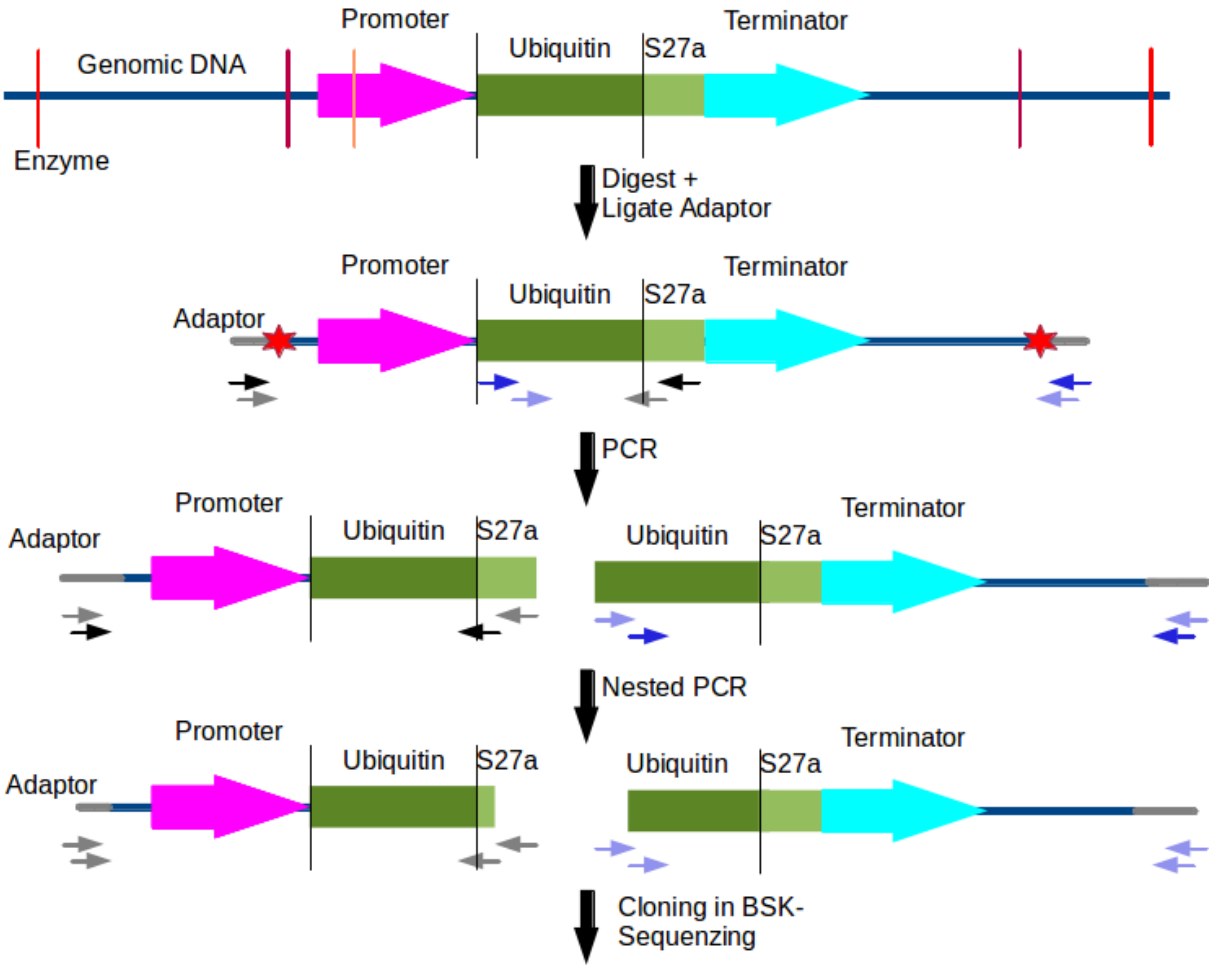


Figure 10.3: Genome walking — schematic illustration of the carried out experiment. Three different blunt-end cutters were used separately to digest the genomic DNA. Red stars indicate the ligation of the adaptors. Two separate, consecutive PCR reactions were carried out with nested primers. The PCR product from the first reaction was diluted 1:100 and used as a template for the nested PCR reaction. The black primer set is amplifying the promoter region, and the blue primer set the terminator region. Primers not involved in the reaction are indicated transparent for guidance.

Table 10.8: Genome walking - PCR program and component concentrations used to amplify an endogenous ubiquitin promoter/terminator sequence of the unsequenced *C. cuspidata* genome. The digested adaptor-ligated genomic DNA fragments (DL-DNA) were added as a template, and for a negative control genomic DNA, as well as the adaptor-ligated genomic DNA (L-DNA) were used. In the first PCR reaction the Primer pair was AP1 with the corresponding ubiquitin, respectively S27a primer (see 10.2). The second PCR reaction was carried out with the nested AP2 primer and the nested Ubiquitin primers (for schematic illustration see 10.3, for the primer see 10.2). In total the second PCR was carried out with 50 different templates and primer combinations. The PCR setup is based on the Genome Walker^{TM} Human Kit.

Component	Amount in μl	**Cycle - Temperature**	Time
DL-DNA	1	1) 98 °C	2 min
Buffer	2	2) 95 °C	30 sec
Phusion®	0.1	3) 61 °C	20 sec
dNTP	0.2	4) 72 °C	3 min
AP1	0.5	Repeat cycle 2-4 for 7 times	
Prim	0.5	5) 95 °C	30 sec
MQ	5.7	6) 61 °C	20 sec
		7) 67 °C	3 min
		8) 67 °C	7 min
		Repeat cycle 5-7 for 32 times	
Dilute PCR product 1:100 with MQ			
Diluted PCR product	1	1) 98 °C	2 min
Buffer	2	2) 96 °C	30 sec
Phusion®	0.1	3) 94 °C	20 sec
dNTP	0.2	4) 72 °C	3 min
AP2	0.5	Repeat cycle 2-4 for 5 times	
Nested Primer	0.5	5) 96 °C	30 sec
MQ	5.7	6) 67 °C	20 sec
		7) 72 °C	3 min
		8) 67 °C	7 min
		Repeat cycle 5-7 for 22 times	

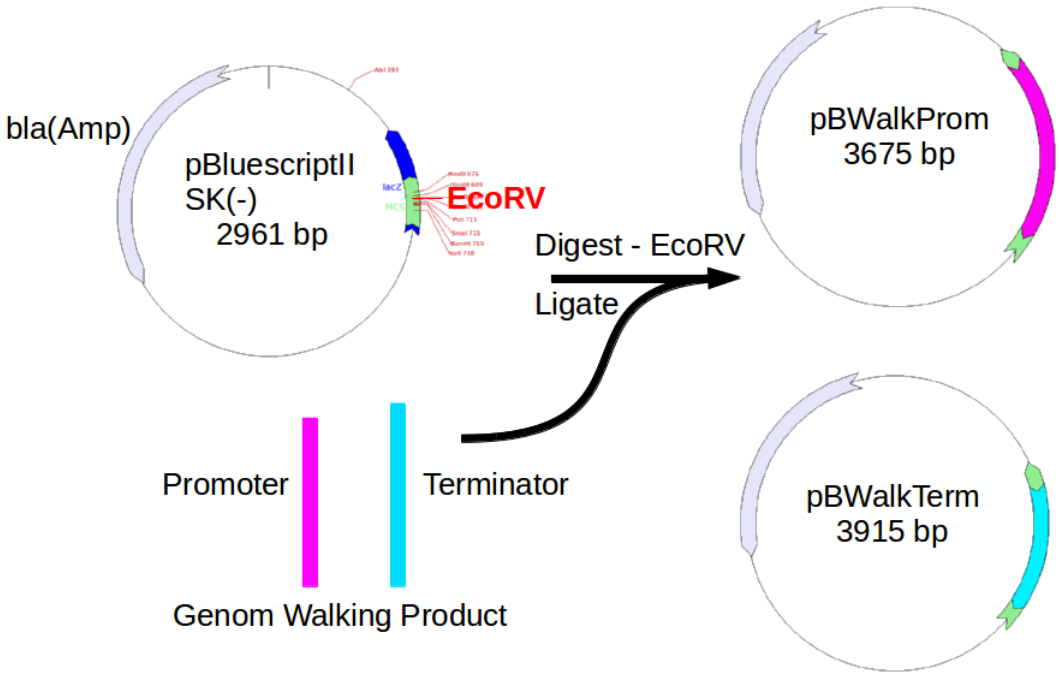


Figure 10.4: Schematic illustration of the performed genome walking products cloned into the BSK- vector. Pink — promoter; turquoise — terminator; green — multiple cloning site; blue — lacZ gene; light gray — ampicillin resistance gene (bla).

Table 10.9: Endogenous C. cuspidata ubiquitin promoter/terminator sequences used for cloning the expression cassette. Black — terminator/promoter; red — ubiquitin sequence; blue — intron.

Promoter - long 583 bp
CTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCCTGTGCAGGACCTATAAATAAGCGTAAACACG
ATTGCGAAAATCCTCCAAACAGCAGCGAGCGAGCGTGCTGGCTGGGCGGAGGGCTCTAGGGACCTAC
TACACTCACTAGGGCCTATGATACACGGAACGGACCTCGTAGGAAATTCCCAGAATTCGCGTGACTC
CGCGTAGTTCTCAAACGCGATTTGGTAGGGCCCTACCAAGCAGGTACAGTACCATGTGTTAACGACT
GGATCGCCATTGCATCATCAACGGCAGTACCTGGTAGTCTCAAGCTTCCTCTGATTATCTTGCTATAAA
CTCGCAGAAGCGAAATTTTCATTCAGCAACGACAACAAACAATGCAGCAGAAATTGATCTGTGTAGTG
AGAACATAGTGGTTCCCCGATAACATGTTTCATATTAACAATGAAATCCACGCGGCATGTTTGATGCGG
AAAGTAGCCATGACAGAACTTTTATTTCTTCAGCTAATAGTTAACGTCTCTTTTTTTTTCAATTCTAATAG
Promoter - short 507 bp
CTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCCTGTGCAGGACCTATAAATAAGCGTAAACACGA
TTGCGAAAATCCTCCAAACAGCAGCGAGCGAGCGTGCTGGCTGGGCGGAGGGCTCTAGGGACCTACTAC
ACTCACTAGGGCCTATGATACACGGAACGGACCTCGTAGGAAATTCCCAGAATTCGCGTGACTCCGCGT
AGTTCTCAAACGCGATTTGGTAGGGCCCTACCAAGCAGGTACAGTACCATGTGTTAACGACTGGATCGC
CATTGCATCATCAACGGCAGTACCTGGTAGTCTCAAGCTTCCTCTGATTATCTTGCTATAAACTCGCAG
AAGCGAAATTTTCATTCAGCAACGACAACAAACAATGCAGCAGAAATTGATCTGTGTAGTGAGAACAT
AGTGGTTCCCCGATAACATGTTTCATATTAACAATGAAATCCACGCGGCATGTTTG
Terminator 489 bp
GTGATTGTAAGCCGCTGAAAGTCGTGGCAATCAATTAGCAGAACCTAGTTATGCTTTGATTCATTTAAA
TCCTACTATCAGATCTTACTACTATTAGATGATTGTCTGCTTTGAATGTTCTACACACCACGAACTGTG
CAAAGAGGAACATCCAACACTACAACCTGTGCAGTGAGGATGCTTGTAACTCTGTGAAAAAGTGTTAGG
TCCAGTTTTTGTACGTTGCAATCCCAAATCTTTGCATGTATTGTAAGTAAGCTGCCCTAACGTTAGTTC
ACTAAAGCGTCTCCATACGTGTCTCGGCATCTGGCTAATGCCCAAATTGGAAAAACATTGCGATAGGCG
GTGTTTGTGATCCCACATGCACCTGTAAACGCCCCAGCAAATCCCTTCTTGAGGCCAATCGCCACAAGG
GAGCTGCCTGCTAATGAGATACTGCACGCCCATTGTTACAGCTTTCACATCTTCACACTTTGCTAAGGA

of the promoter sequence (507 bp) were cloned into the EcoRV recognition site of the pBT vector (Fig. 10.5, Tab. 10.9). All the following steps were carried out for both, the vector with the long promoter pBlPT and the short promoter version PBsPT, respectively. The revers primers for PCR-amplification of the promoters were designed in a way, that both resulting promoter versions had an 3' extension with a StuI recognition site. This site was chosen for further cloning of selectable markers / reporter genes (Fig. 10.5). Unfortunately, the StuI recognition site resulted to be methylated. To be able to use the StuI recognition site for further cloning the methylation was PCR-point mutated (for corresponding primer see 10.2) and an additional EcoRV recognition site was simultaneously inserted (Tab. 10.10).

Table 10.10: Mutation - PCR program and component concentrations used to point mutate the methylated StuI recognition site and simultaneously insert a EcoRV recognition site into pBlPT and pBsPT.

Component	Amount in μl	**Cycle - parameters**	Time
Vector	0.5	1) 98 ℃	1 min
Buffer	10	2) 98 ℃	10 sec
Phusion®	0.5	3) 55 ℃	30 sec
dNTP	1	4) 68 ℃	30 sec
Prim For	2.5	Repeat cycle 2-4 for 18 times	
Prim Rev	2.5		
MQ	5.2		

The selectable markers natII (ClonNat resistance), neo (G418 resistance) and eGFP (fluorescence reporter gene), were amplified from the following vectors: Neo — pPHA-T1-nptII; NatII — P2-2 and eGFP — 493. The amplified genes were cloned into the inserted EcoRV cutting site, between the ubiquitin promoter and terminator. The resulting vectors pBsPT-Neo, pBlPT-Neo, pBsPT-NatII, pBlPT-NatII, pBsPT-eGFP and pBlPT-eGFP were sent for sequencing and used further on to transform *C. cuspidata*. All intermediate vectors as well as the final vectors were stored at -20 ℃.

Microprojectile bombardment of *C. cuspidata* with plasmids containing an endogenous promoter/terminator

C. cuspidata was transformed by microprojectile bombardment as described above (chapter 10.5.1), with the same experimental settings and the same positive control. The transformation was repeated at least 16 times for each of the four constructed plasmids containing selectable markers and once carried out for the plasmids with the reporter construct eGFP.

10.6 Transformation of *C. fusiformis* to visualize the actin cytoskeleton

10.6.1 Construction of a plasmid containing LA-GFP

The Lifeact (LA) gene fused to eGFP was chosen to label the actin cytoskeleton of *C. fusiformis*. A plasmid was constructed containing two expression cassettes each under the endogenous FCP promoter and terminator of *C. fusiformis*: the first drives the expression of LA-eGFP and the second drives the expression of the zeocin resistance gene ble.

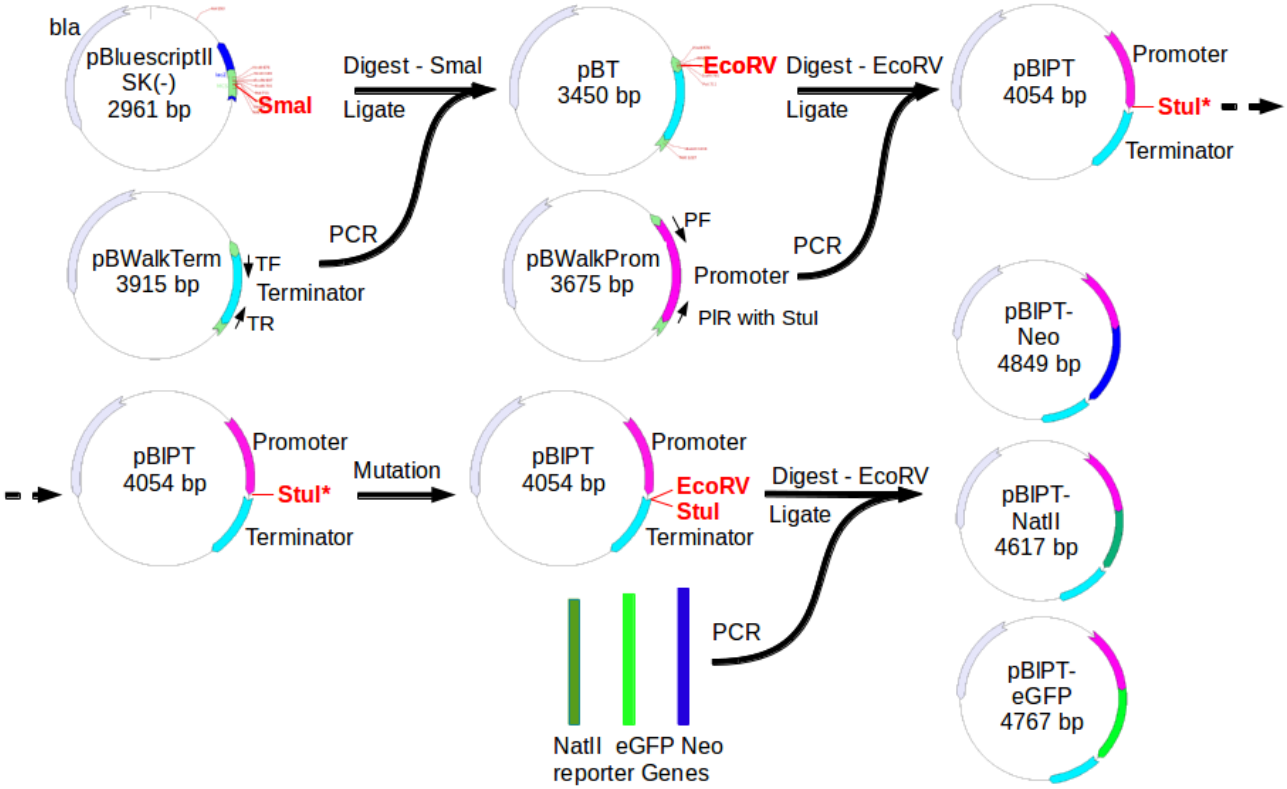


Figure 10.5: Schematic illustration of the performed cloning steps to construct the expression cassette with the endogenous *C. cuspidata* ubiquitin promoter/terminator. All the following steps were carried out for both, the vector with the long promoter pBsPT and the short promoter PBsPT, but are illustrated only for the long promoter. The plasmids pBIPT-Neo, pBIPT-NatII and pBIPT-eGFP, as well as pBsPT-Neo, pBsPT-NatII and pBsPT-eGFP were used to transform *C. cuspidata*. Pink — promoter; turquoise — terminator; light green — eGFP; blue — Neo G418 resistance gene; green — NatII ClonNat resistance gene; light gray — ampicillin resistance gene (bla); red star — methylated StuI cutting site.

Two single-stranded LA sequences LAF and LAR containing the EcoRV and KpnI recognition site were designed and annealed to a double-stranded sequence as described in chapter 10.5.2. All cloning steps of the expression cassette construction are described in detail in Fig. 10.6.

A control vector containing only the eGFP gene and not the LA sequence was also constructed. The PCR-amplified eGFP sequence with the primers GFPEcoRVF and GFPNotIR, was EcoRV-NotI digested and cloned into the EcoRV-NotI digested 333 vector. The following steps were identical to the cloning steps described in Fig. 10.6.

10.6.2 Microprojectile bombardment of *C. fusiformis* to label actin

C. fusiformis was transformed by microprojectile bombardment, as described by Poulsen and Kröger (2005). *C. fusiformis* was shoot four times with the plasmids pPLAGFPT-PbleT (CfLA-GFP) and twice with the plasmid pPGFPT-PbleT (CfGFP).

10.7 Microscopy

For microscopy the cells of *C. fusiformis* were immobilized on 1% polyethylenimine (PEI) coated slides. The microscope, the camera and the software used for documentation are described in chapter 6.3. eGFP was excited at 488 nm and detected between 500-530 nm.

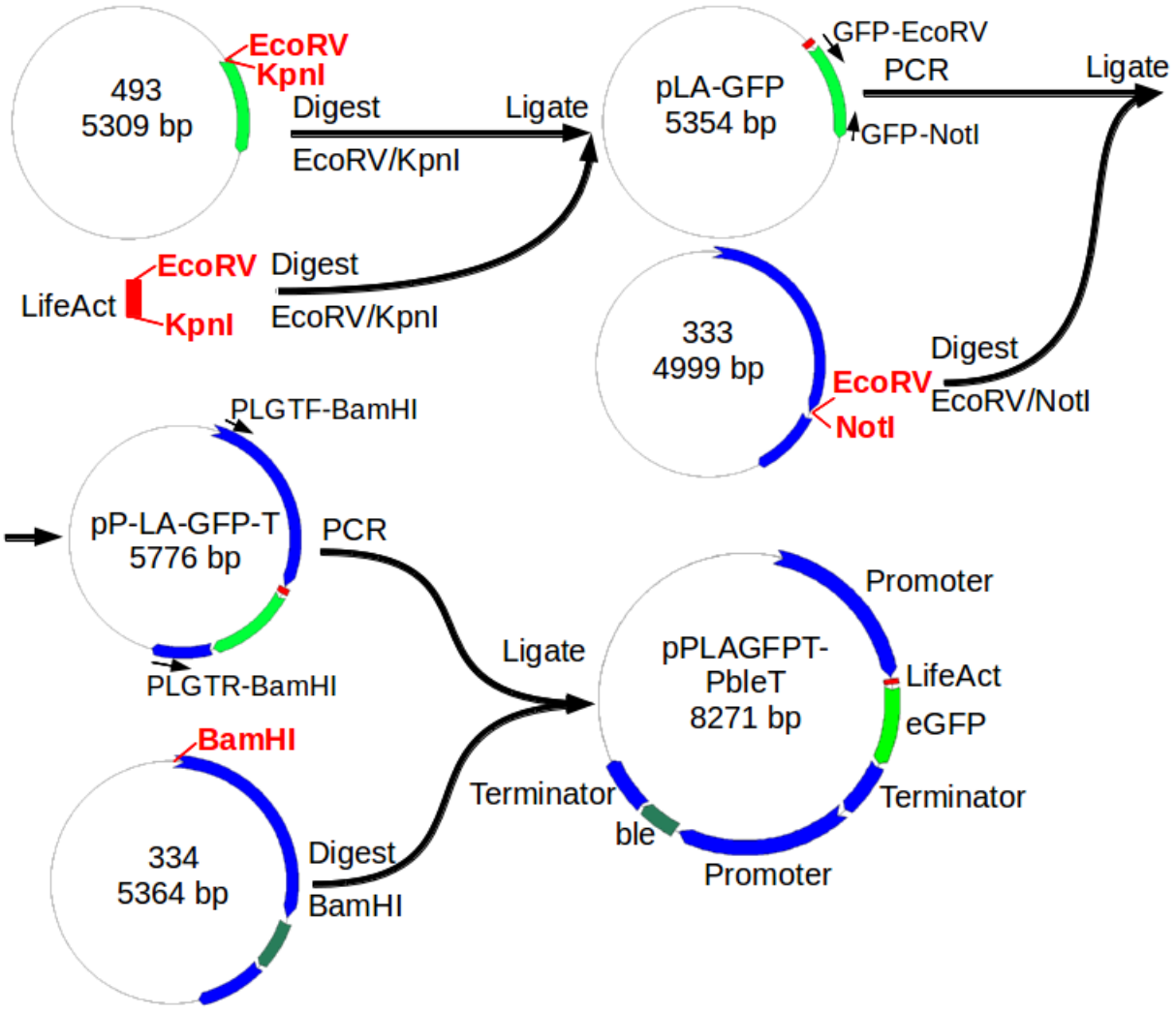


Figure 10.6: Schematic illustration of the cloning steps performed to construct a plasmid with two expression cassettes. The first step was the fusion of LA to eGFP, which was then cloned into the expression cassette of the 333 vector with an endogenous *C. fusiformis* FCP promoter and terminator. The full LA-eGFP expression cassette was then subcloned into the 334 vector upstream of the already existing selection resistance cassette. The resulting plasmid CfLA-GFP contains both expression cassettes. Prior to each ligation the restriction enzymes from the digestion were inhibited. The restriction enzymes and primers used in the PCR reactions are indicated. Blue — FCP-promoter/terminator of *C. fusiformis*; light green — eGFP; red — Lifeact; dark green — ble Zeocin resistance.

11 Results

11.1 Transformation of *C. cuspidata*

The essential prerequisites for transformation were all fulfilled by *C. cuspidata*. The cells grew in colonies on agar plates, antibiotics for selection were chosen (chapter 4.3) and the low pressure of 28 psi did not effected the growth behavior of *C. cuspidata*. Transformation by particle bombardment is known to lead to a devastation in the middle of the sample and to transformants in a doughnut like shape around it. *C. cuspidata* cells in the center of the petri dish, hit by multiple tungsten particles (1100 psi), were perforated and leaky, whereas the cells further outward showed a good morphology.

11.1.1 Transformation of *C. cuspidata* with established diatom transformation plasmids

Plasmids that already have been successfully used by other groups to transform *P. tricornutum*, *T. pseudonana* and *C. fusiformis* were used to transform *C. cuspidata* (Tab. 10.1). Unfortunately none of the four tested vectors (353, 370, P2-2, pPHA-T1-nptII) resulted in transformed *C. cuspidata* cells. Increasing the acceleration pressure from 1100 psi to 1550 psi did not lead to transformation success. Also modifying the recovery phase after transformation by: a) leaving *C. cuspidata* on the agar plate for up to three days, b) adding liquid medium onto the agar plate, or c) even transferring the cells into liquid medium for recovery, before transferring them onto the plates containing the antibiotics did not result in any transformants.

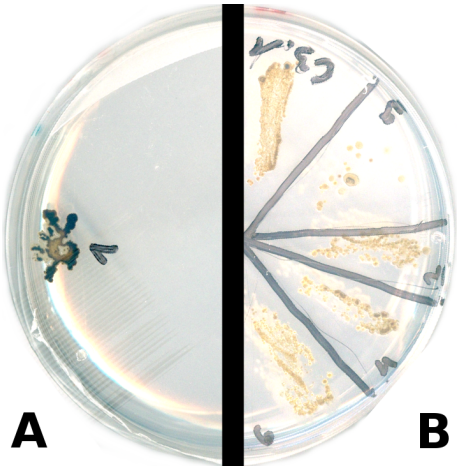


Figure 11.1: *P. tricornutum* was transformed by microprojectile bombardment. **A:** Single colony of a transformant. The colony is already a few months old and therefore very large. Most plates showed 2-5 colonies. **B:** Storing five different transformants on a zeocin-agar plate.

The positive control, transformation of *P. tricornutum* with the StuI-GFP-pPhaT1, resulted in up to 15 colonies per shoot (Fig. 11.1). Colonies appeared 8 to 14 days after transformation.

11.1.2 Transformation of *C. cuspidata* with an endogenous promoter/terminator

DNA Extraction

Table 11.1: Sequence of the two C. cuspidata actin genes ActA and ActB.

ActA

CTATAGAGGCTCACCCTGTTCTCCTCACGGAAGCTCCCCAAAATCCTAAGGCTAACCGCGAGCGCATGA
CACAAATCATGTTCGAGACTTTCAACGTCCCCGCCATGTATGTGAACATTCAGGCAGTTTTGTCTCTGT
ACGCTTCGGGTCGTACTACTGGATGCGTGTTGGACTCTGGTGATGGTGTAACCCACACTGTGCCGATTT
ACGAAGGTTACGCCCTTCCTCATGCCGTCATTCGTTTGGATCTTGCTGGACGTGACTTGACCGATTACT
TGATGAAGATTCTCACGGAACGTGGTTACTCTTTGACCACCACGGCAGAGCGTGAAATTGTTCGAGATA
TCAAGGAATCGCTTTGCTTCGTTGCAGTAGACTTTGAGGAGGAAATGAAAAAGGCAGCTGAATCTTCGG
CTCTTGAGAAGTCTTTCGAACTTCCCGATGGTAACATCATTGTGATTGGAAACGAACGCTTCCGTT
GCCCGGAGGTCTCTTTTC

ActB

CAGTTTTGTTGACCGAAGCCCCAATGAACCCTAAATCTAACCGTGAAAAGATGACCCAAATCATGTTTG
AAACCTTCAACGTACCAGCCTTCTACGTTTCTATCCAAGCCGTCTTATCTTTATACTCGTCTGGTAGAA
CCACCGGTATCGTTTTAGATTCGGGTGATGGTGTTACTCACGTTGTTCCAATTTACGCCGGTTTCTCCT
TACCACACGGTATCTTAAGAATTGACTTGGCTGGTAGAGACTTGACCGACTATTTGATGAAGATCTTGT
CCGAAAGAGGTTACACTTTCTCCACCACCGCTGAAAGAGAAATTGTCCGTGATATCAAAGAAAAATTAT
GTTACGTTGCCTTGGACTTTGAACAAGAAATGCAAACTTCATCTCAATCCTCCGCCATCGAAAAGTCTT
ACGAATTACCTGATGGCCAAGTTATTACTATCGGTAACGAAAGATTTAGAGCTTCCGAAGCTTTGTTCC
GTCCTTCTGACTTAGGTTTAGAAGCCGCTGGTATTGACCAAACCACTTACAACTCGATTATGAAGTGTG
ATGTCGATGTCAGAAAGGAATTATACGGTAACATTGTTATGTCTGGTGGTACCCCATT

To amplify genomic DNA fragments with degenerated primers, high quality DNA is needed. The modified CTAB protocol was suitable for the extraction of genomic DNA at the appropriate quality (Fig. 11.2).

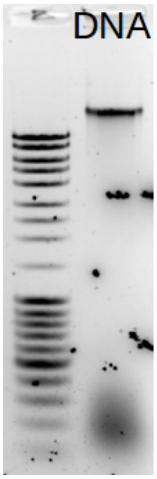


Figure 11.2: CTAB DNA extraction. Mass ruler.

C. cuspidata actin gene

Two genomic DNA fragments were successfully amplified with degenerate primers designed for actin and resulted in two actin gene fragments *ActA* and *ActB* (Tab. 11.1). Much effort was put into amplifying the promoter and terminator region of an actin gene by inverse PCR, but none of the approaches gave a satisfying result. Steps and intermediate data of this fruitless endeavor are not covered here, but listed in the appendix.

C. cuspidata ubiquitin gene

The EST databases of *P. tricornutum* and *T. pseudonana* were screened for ubiquitin. Both organisms showed a high and constant ubiquitin expression under various culturing conditions.

The degenerated primer pairs UF7-UR6, UF7-UR7, UF8-UR6, UF8-UR7, UF9-UR7 all amplified sequences resulting in strong and specific DNA fragments at the expected sizes, although the primers have been highly degenerated (UF7 is a mixture of 768 different primers)(Fig. 11.3). Two genomic DNA fragments were successfully amplified differing slightly in size (Fig. 11.3). These fragments were cloned into the BSK-vector resulting in many white colonies after blue-white selection (Fig. 11.4); 4 out of 12 white colonies had the PCR product inserted. The two sequenced gene fragments were coding for ubiquitin and had a sequence identity of 85% on the nucleotide level (Tab. 11.2, Fig. 11.6). The two fragments code for the same protein, as all differences on the

nucleotide level were conserved leading to a 100% identity on amino acid level.

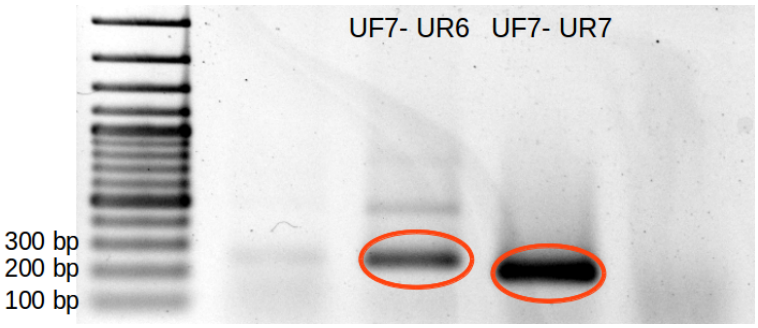


Figure 11.3: Amplified ubiquitin gene by degenerated primers. Red encircled bands were cloned into the BSK- vector and sent for sequencing.

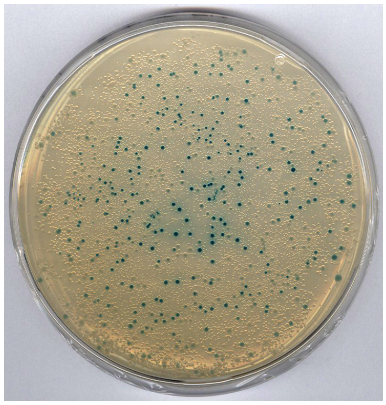


Figure 11.4: Blue-white selection of the resulted PCR product (UF7-UR7) cloned into the BSK- vector.

Table 11.2: Sequence of the two C. cuspidata ubiquitin-S27a genes Ubi1 and Ubi2. Primer pairs which amplified the genes are listed.

Ubiquitin 1 (UF7-UR7)
ATGCAGATCTTCGTGAAAACCCTGACAGGAAAGACGATCACGCTCGATGTTGAGCCGTCTGACACCAT
TGACAACGTCAAAGCCAAGATTCAGGATAAGGAGGGTATCCCCCCTGATCAACAGCGTTTGATTTTCG
CCGGAAAGCAGTTGGAAGATGGACGCACTCTCTCGGACTACAACATCCAGAAAGAATCCACCCTTC
Ubiquitin 2 (UF7-UR6)
ATGCAGATCTTCGTCAAGACCTTGACTGGAAAGACAATTACCTTGGATGTCGAGCCTTCTGACACCAT
CGACAACGTCAAGGCAAAGATTCAAGATAAGGAGGGAATCCCCCCTGACCAACAACGATTGATTTTTG
CTGGTAAACAGTTGGAAGATGGACGCACTTTGTCGGACTACAACATTCAAAAGGAGTCTACTCTTCAC
TTGGTCCTTCGTCTGCGAGGTGGGGGCAAGAA

Endogenous ubiquitin promoter/terminator of *C. cuspidata*

The amplification of the promoter/terminator sequence up- and downstream of the two ubiquitin genes by genome walking resulted in bands between 500 bp and 1000 bp (Fig. 11.5). The band at 714 bp was sequenced and consisted of a ubiquitin gene (*Ubi3*) fragment (335 bp) and an upstream promoter region (379 bp). The ubiquitin gene had a 160 bp long intron shortly after the start codon (Tab. 11.4). The second amplified genome walking PCR product was 954 bp long containing a 234 bp long full coding sequence of a ubiquitin gene (*Ubi4*) and downstream of the ubiquitin gene a 720 bp long terminator sequence (Tab. 11.4). The two amplified ubiquitin genes *Ubi3* and *Ubi4* differ from the former two sequenced ubiquitin sequences *Ubi1* and *Ubi2* (Fig. 11.6). Resulting in a total of four different ubiquitin gene sequences amplified in the present work, as a by-product.

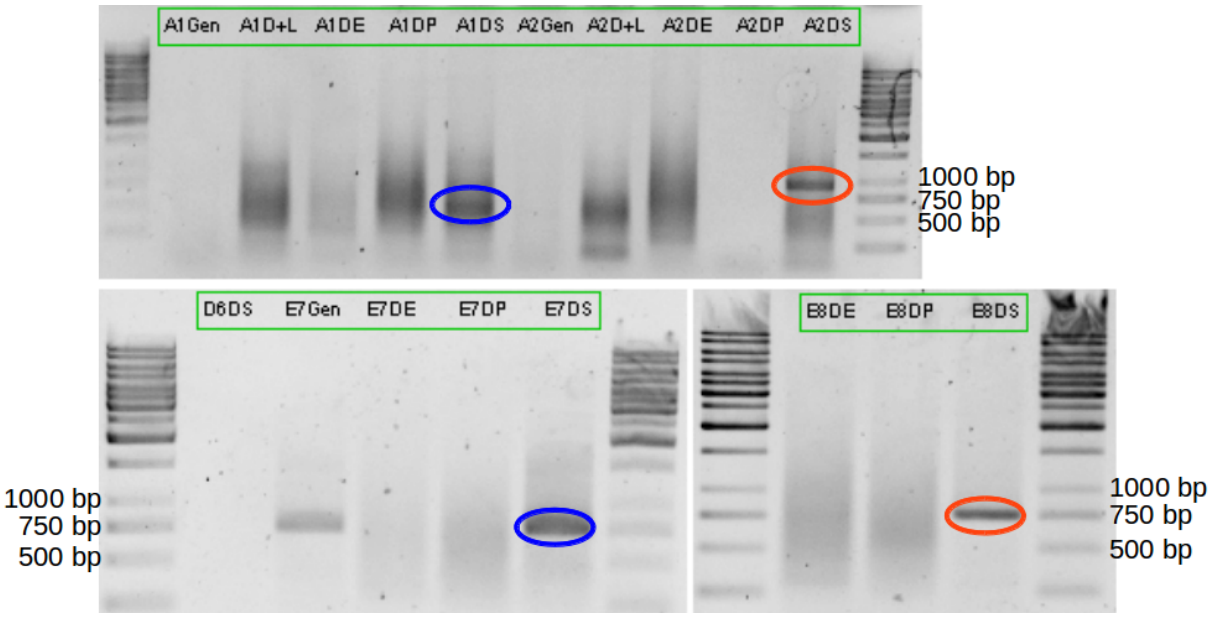


Figure 11.5: Selected results of the ubiquitin promoter/terminator amplification by genome walking. The labeling of the samples is explained in (Tab. 11.3) Blue and red marked bands were sent for sequencing. The sequences of the bands encircled in red were selected to further clone the expression cassette designated to transform *C. cuspidata*.

```
Ubi3  ATGCAGCAGAAATTGATCTGTGTAGTGAGAACATAGTGGTTCCCCGATAACATGTTTCATATTAACAATGAAATCCACGCGGCATGTTTGATGCGGAAA
Ubi1  ATGCAG----------------------------------------------------------------------------------------------
Ubi2  ATGCAG----------------------------------------------------------------------------------------------
Ubi4  ATGCAG----------------------------------------------------------------------------------------------
      ******

Ubi3  GTAGCCATGACAGAACTTTTATTTCTTCAGCTAATAGTTAACGTCTCTTTTTTTTCAATTCTAATAGATTTTTGTGAAGACCTTGACGGGAAAGACGAT
Ubi1  ------------------------------------------------------------------ATCTTCGTGAAAACCCTGACAGGAAAGACGAT
Ubi2  ------------------------------------------------------------------ATCTTCGTCAAGACCTTGACTGGAAAGACAAT
Ubi4  ------------------------------------------------------------------ATCTTCGTGAAAACCCTGACAGGAAAGACGAT
                                                                         ** ** ** **.*** **** ********.**

Ubi3  CACATTGGATGTCGAGCCTTCTGACACCATTGACAACGTCAAGGCAAAGATTCAGGATAAGGAGGGAATTCCCCCTGACCAACAACGATTGATTTTTGC
Ubi1  CACGCTCGATGTTGAGCCGTCTGACACCATTGACAACGTCAAAGCCAAGATTCAGGATAAGGAGGGTATCCCCCCTGATCAACAGCGTTTGATTTTCGC
Ubi2  TACCTTGGATGTCGAGCCTTCTGACACCATCGACAACGTCAAGGCAAAGATTCAAGATAAGGAGGGAATCCCCCCTGACCAACAACGATTGATTTTTGC
Ubi4  CACGCTCGATGTTGAGCCGTCTGACACCATCGACAACGTCAAAGCCAAGATTCAGGATAAGGAGGGTATCCCCCCTGATCAACAGCGTTTGATTTTCGC
       **  * ***** ***** ********** **********.**.********.**********:** ******** *****.**:******** **

Ubi3  TGGTAAACAGTTGGAAGATGGACGCACTTTGTCGGACT-----------------------------------------------------------
Ubi1  CGGAAAGCAGTTGGAAGATGGACGCACTCTCTCGGACTACAACATCCAGAAAGAATCCACCCTTC-------------------------------
Ubi2  TGGTAAACAGTTGGAAGATGGACGCACTTTGTCGGACTACAACATTCAAAAGGAGTCTACTCTTCACTTGGTCCTTCGTCTGCGAGGTGGGGGCAAGAA
Ubi4  TGGAAAGCAGTTGGAAGATGGACGCACTCTCTCCGACTACAACATTCAGAAGGAATCTACTCTTCATTTGGTCCTTCGTCTTCGTGGAGGAAATTAA--
       **:**.********************* * ** ****
```

Figure 11.6: Aligned *C. cuspidata* ubiquitin gene sequences. Created with the alignment program Kalign from the EMBL-EBI website.

Table 11.3: Explanation of the genome walking sample labeling. Exemplification: the successful amplified terminator sequence E8DS was digested by StuI and amplified with the primer set AP1 — Ub2P1.2 followed by the nested primer set AP2 — Ub2P1.2

Gen	Genomic DNA (negative control)		
D+L	Genomic DNA adaptor ligated (negative control)		
DE	Gen. DNA digested EcoRV, adaptor ligated		
DP	Gen. DNA digested PvuII, adaptor ligated		
DS	Gen. DNA digested StuI, adaptor ligated		
First primer set		Second primer set (nested)	
1	AP1—Ub1T1.1	A	AP2—Ub1T1.2 for 1 and 2
2	AP1—Ub1T2.1	B	AP2—Ub1T2.2 for 1 and 2
3	AP1—Ub1P1.1	C	AP2—Ub1P1.3 for 3 and 4
4	AP1—Ub1P2.1	D	AP2—Ub2T1.2 for 5 and 6
5	AP1—Ub2T1.1	E	AP2—Ub2P1.2 for 7 and 8
6	AP1—Ub2T2.1		
7	AP1—Ub2P1.1		
8	AP1—Ub2P1.2		

Table 11.4: C. cuspidata ubiquitin promoter/terminator sequence. The Ubiqutin3 sequence is 714 bp long containing a 379 long promoter fragment and a 335 bp long 5'-coding sequence of a ubiquitin gene, including a 160 bp long intron. The Ubiquitin4 sequence is 954 bp long containing the 234 bp long complete reading frame of a ubiquitin gene and a 720 downstream terminator region. Red — ubiquitin sequence; blue — intron; black — terminator/promoter

Promoter - Ubiquitin3

CTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCCTGTGCAGGACCTATAAATAAGCGTAAACACG
ATTGCGAAAATCCTCCAAACAGCAGCGAGCGAGCGTGCTGGCTGGGCGGAGGGCTCTAGGGACCTAC
TACACTCACTAGGGCCTATGATACACGGAACGGACCTCGTAGGAAATTCCCAGAATTCGCGTGACTC
CGCGTAGTTCTCAAACGCGATTTGGTAGGGCCCTACCAAGCAGGTACAGTACCATGTGTTAACGACT
GGATCGCCATTGCATCATCAACGGCAGTACCTGGTAGTCTCAAGCTTCCTCTGATTATCTTGCTATAAA
CTCGCAGAAGCGAAATTTTCATTCAGCAACGACAACAAACAATGCAGCAGAAATTGATCTGTGTAGTG
AGAACATAGTGGTTCCCCGATAACATGTTTCATATTAACAATGAAATCCACGCGGCATGTTTGATGCGG
AAAGTAGCCATGACAGAACTTTTATTTCTTCAGCTAATAGTTAACGTCTCTTTTTTTTTCAATTCTAATAG
CATGACAGAACTTTTATTTCTTCAGCTAATAGTTAACGTCTCTTTTTTTTTCAATTCTAATAGATTTTTG
TGAAGACCTTGACGGGAAAGACGATCACATTGGATGTCGAGCCTTCTGACACCATTGACAACGTCAAGG
CAAAGATTCAGGATAAGGAGGGAATTCCCCCTGACCAACAACGATTGATTTTTGCTGGTAAACAGTTGG
AAGATGGACGCACTTTGTCGGACT

Terminator - Ubiquitin4

ATGCAGATCTTCGTGAAAACCCTGACAGGAAAGACGATCACGCTCGATGTTGAGCCGTCTGACACCATC
GACAACGTCAAAGCCAAGATTCAGGATAAGGAGGGTATCCCCCCTGATCAACAGCGTTTGATTTTCGCT
GGAAAGCAGTTGGAAGATGGACGCACTCTCTCCGACTACAACATTCAGAAGGAATCTACTCTTCATTTGG
TCCTTCGTCTTCGTGGAGGAAATTAAGTGATTGTAAGCCGCTGAAAGTCGTGGCAATCAATTAGCAGAA
CCTAGTTATGCTTTGATTCATTTAAATCCTACTATCAGATCTTACTACTATTAGATGATTGTCTGCTTT
GAATGTTCTACACACCACGAACTGTGCAAAGAGGAACATCCAACACTACAACCTGTGCAGTGAGGATGC
TTGTAACTCTGTGAAAAAGTGTTAGGTCCAGTTTTTGTACGTTGCAATCCCAAATCTTTGCATGTATTG
TAAGTAAGCTGCCCTAACGTTAGTTCACTAAAGCGTCTCCATACGTGTCTCGGCATCTGGCTAATGCCCA
AATTGGAAAAACATTGCGATAGGCGGTGTTTGTGATCCCACATGCACCTGTAAACGCCCCAGCAAATCCC
TTCTTGAGGCCAATCGCCACAAGGGAGCTGCCTGCTAATGAGATACTGCACGCCCATTGTTACAGCTTT
CACATCTTCACACTTTGCTAAGGATAGAGCTAGGAGTGCCCAAGCCGTACTGACAACGCCACTGCCA
CCGTCGCCATAGGCTTCCATCCCATTCTCGGAGTATTTCCTATCATAACAAGATGTTAAATCTTCAC
CCCATCCTCCATTTGGACGCTGATGGCGCAATAGGAACTGACAAGCCTTCTGGACGGCGGGTGATGTC
AGTGGTTCTCCAGACCTCACAAGGACCAGCCCGGGCCGTCGACCACGCGTGCCCTATAG

Plasmid construction

Cloning the endogenous ubiquitin promoter and terminator into the BSK- vector resulted in an expression cassette with a long (pBlPT) and with a short promoter variant (pBsPT). The long promoter contained the 5' region of the Ubiquitin3 gene including the full length of the first intron. The short promoter variant contained the 5' region of the Ubiquitin3 gene including only 85 bp of the intron. Unfortunately, both the pBlPT and pBsPT had a methylated StuI recognition site, which was planed to be the major cloning site for the selection markers and the reporter gene. To overcome the methylation at the StuI recognizing site, the plasmids were point mutated by PCR resulting in an exchange of an A to T. Simultaneously an additional EcoRV recognizing site was successfully inserted upstream of the StuI recognition site by the PCR-based mutation. The two selection markers (neo, natII) as well as the reporter gene (eGFP) were cloned into the EcoRV cutting site resulting in six final plasmids used to transform Cc (Fig. 11.7).

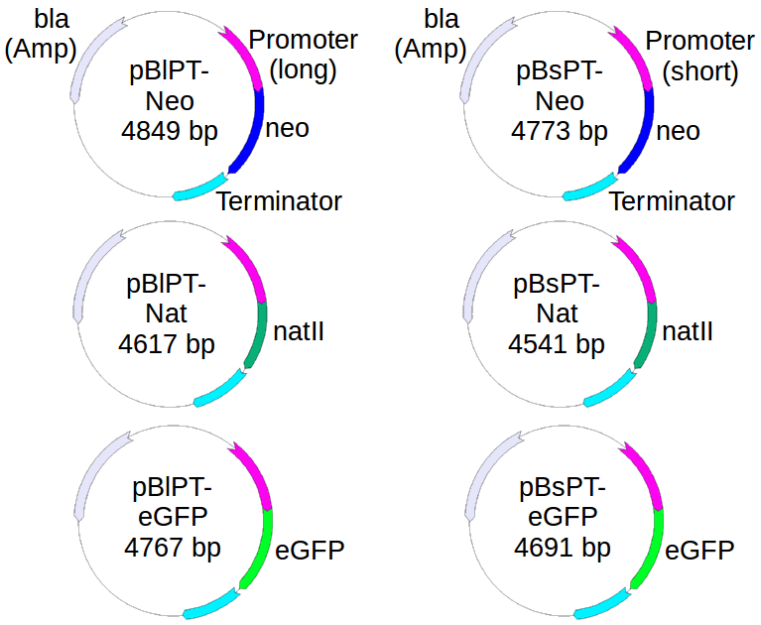


Figure 11.7: Final expression cassettes including reporter genes and selection markers under the control of the endogenous ubiquitin promoter/terminator. A total of six vectors were cloned with three different genes: NatII — ClonNat resistance; Neo — G418 resistance; eGFP — green fluorescent protein

Microprojectile bombardment of *C. cuspidata* with plasmids containing an endogenous promoter/terminator

Attempts to transform *C. cuspidata* by partial bombardment with the six plasmids under various parameters and selection conditions (for detail see chapter 11.1.1) did not result in a single transfected cell.

11.2 Transformation of *C. fusiformis* to visualize the actin cytoskeleton

11.2.1 Construction of a plasmid containing LA-GFP

To visualize the actin cytoskeleton of *C. fusiformis in vivo*, an expression cassette was constructed containing the actin reporter Lifeact (LA) fused to eGFP. Besides this a second plasmid only containing eGFP not fused to the LA sequence was also constructed.

eGFP as well as LA-GFP were successfully cloned into the 333 vector (Fig. 11.8 B). The expression cassette containing the endogenous FCP- promoter (FCPp) and the corresponding terminator (FCPt) enclosing the LA-GFP, respectively just eGFP, were successfully cloned into the 334 vector upstream of the zeocin resistance expressions cassette (Fig. 11.9 B).The sequences of

the two resulting plasmids CfLA-GFP, and CfGFP (Fig. 11.10) were reconfirmed by sequencing.

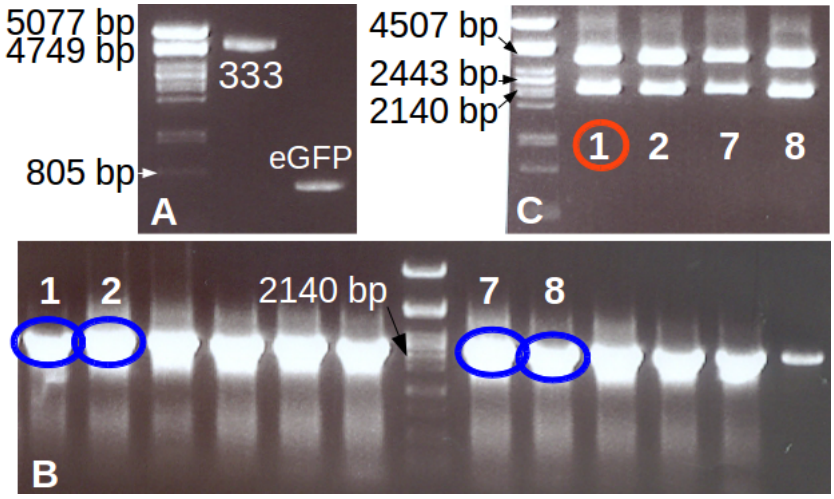


Figure 11.8: Results of cloning steps to construct the expression cassette FCPp-eGFP-FCPt. **A:** The 4988 bp band represents the 333 vector and the 829 bp band is the PCR-amplified eGFP. Both were double digested with EcoRV and NotI. **B:** Colony PCR of 12 colonies (transfected with pP-GFP-T) using the primer pair M13F and SLAR. All 12 colonies were positive. Clones marked by blue circles were further tested by digestion. **C:** Isolated plasmids of the clones 1,2,7 and 8 were tested by EcoRI digestion, all were positive. Clone 1 was used for further cloning.

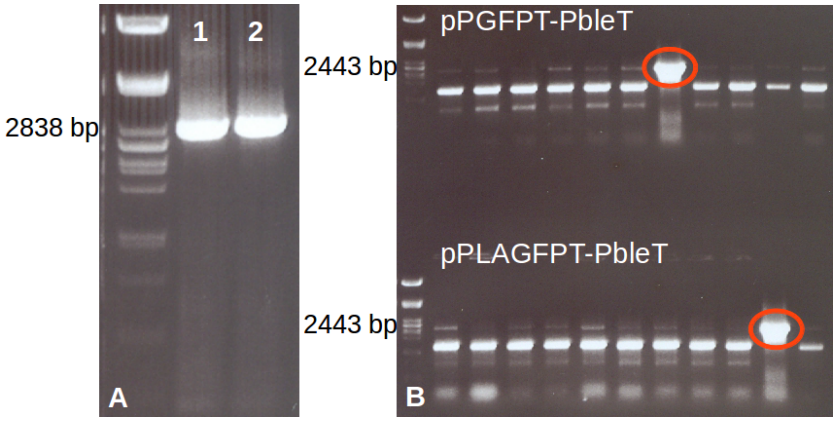


Figure 11.9: Cloning results of the plasmid containing two expression cassettes. **A:** PCR product 1 — promoter/GFP/terminator with 2839 bp; 2 — promoter/LA-GFP/terminator with 2901 bp, both with a 5' and 3' BamHI recognition site. **B:** Colony PCR of 11 colonies for CfLA-GFP and CfGFP, one colony each was positive and sent for sequencing - red circle.

11.2.2 Microprojectile bombardment of *C. fusiformis*

The plasmid CfLA-GFP was shot four times resulting in nine colonies after 10 days (Fig. 11.11). Of these nine zeocin-resistant clones, four colonies were expressing eGFP. Two transformation attempts with the CfGFP plasmid did not result in colonies.

The fluorescent actin reporter LA-GFP was stably transformed in *C. fusiformis*. The observed growth rate, division and locomotion of *C. fusiformis* after the transformation with LA-GFP were just the same as in wild type. Therefore, it can be concluded, that the fusion construct does not interfere with actin dynamics and actin-dependent processes.

The LA-GFP fusion construct labeled thick actin bundles (Fig. 11.12 D) as well as fine actin

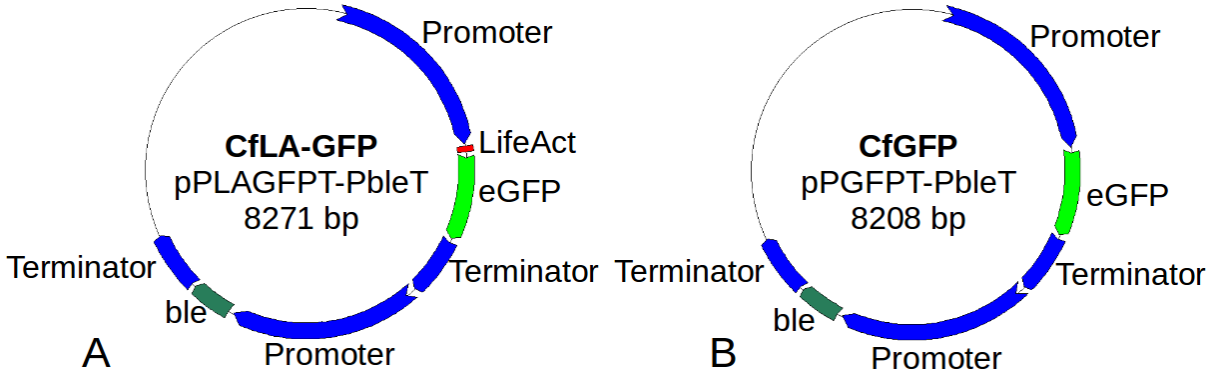


Figure 11.10: Final expression cassettes including the reporter genes under an endogenous promoter/terminator to transform *C. fusiformis* . Lifeact-GFP as an reporter to visualize actin. **A:** Structure of CfLA-GFP. **B:** CfGFP - same vector, but without Lifeact. Promoter — FCP from *C. fusiformis*; Terminator — FCP from *C. fusiformis*; Reporter genes: Ble — Zeocin resistance; eGFP — green fluorescent protein

meshworks (Fig. 11.12 E) in *C. fusiformis*. It seemed from the images that just one thick actin bundle was located underneath each raphe slit instead of two as shown for *C. cuspidata* (chapter 7).

To confirm that Lifeact is labeling F-actin, the transformants were treated with latrunculin A, an F-actin depolymerizing drug. In contrast to *C. cuspidata* (chapter 8) the latrunculin A treatment was efficient in *C. fusiformis*. After 100 min the thick actin bundles mostly disappeared, but filaments were still visible. Within 170 min the depolymerization further advanced, with only a few filaments remaining. After an inhibitor treatment of 210 min actin filaments completely depolymerized. The LA-GFP pattern therefore clearly changed over time under the treatment of latrunculin A (Fig. 11.13).

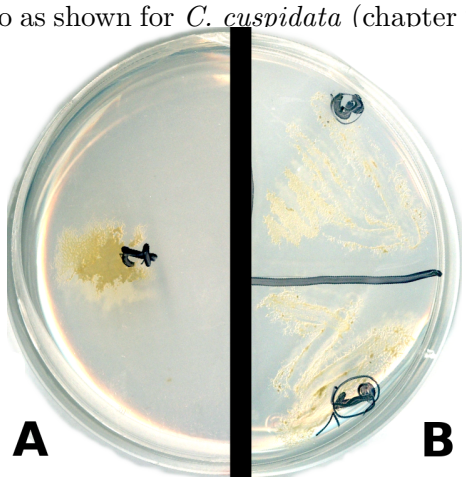


Figure 11.11: Transformed *C. fusiformis* cells by microprojectile bombardment. **A:** Single colony of a transformant expressing CfLA-GFP. The colony is a few months old and therefore very large. **B:** Two transformants stored on a zeocin-agar plate.

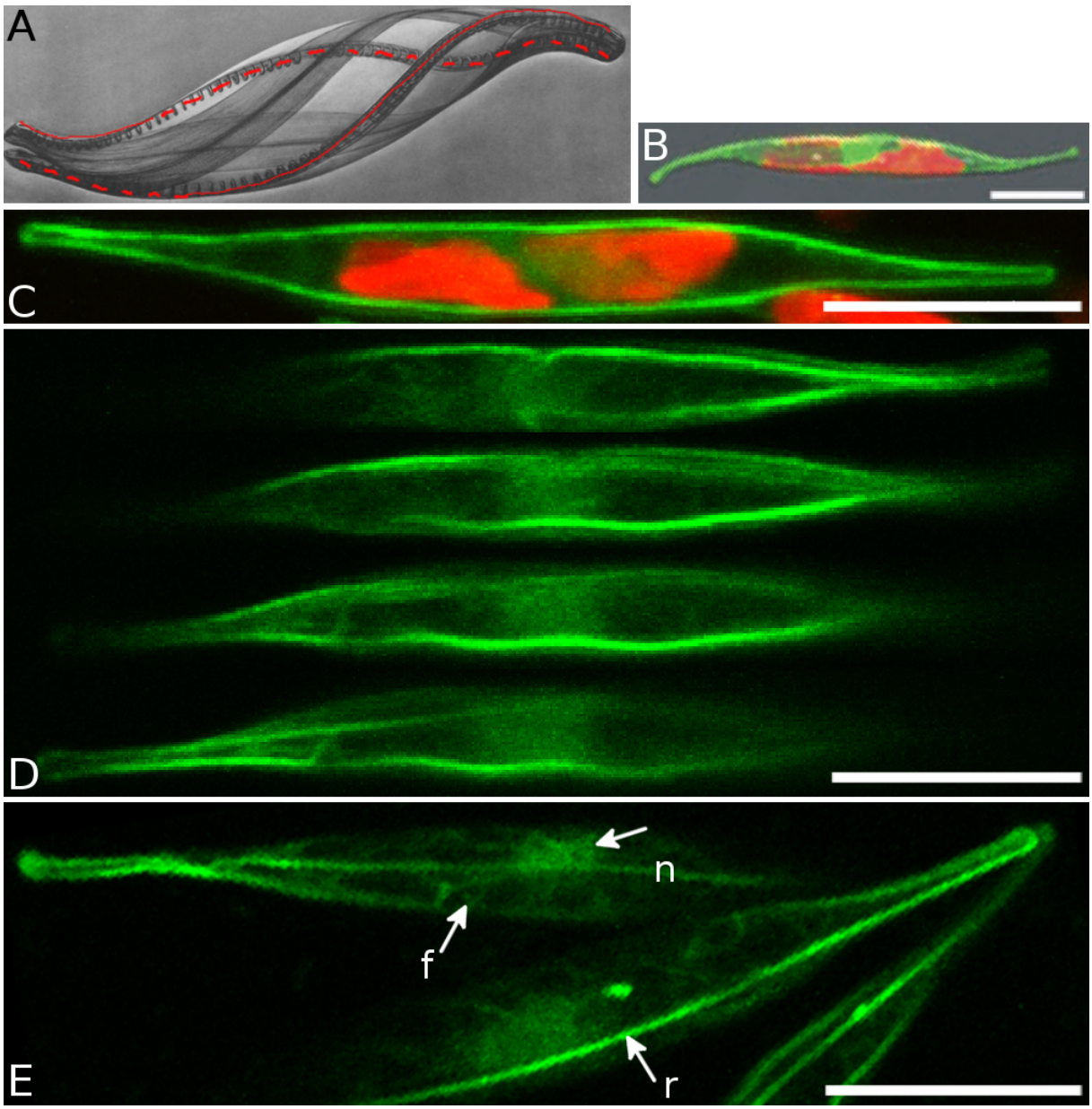


Figure 11.12: Transformed *C. fusiformis* with LA-GFP. **A:** Frustule of *C. fusiformis*, the twisted raphe is marked in red. The raphe facing the observer is drawn as a solid red line, the curved raphe part facing the back of the cell is indicated as a stippled red line (modified from Reimann et al., 1965). Note the regular fibulae spanning over the raphe. **B:** Appearance of a *C. fusiformis* cell transformed with just the eGFP, image taken from Poulsen and Kröger (2005) for comparison. **C:** Full cell projection of all optical sections. *C. fusiformis* cell transformed with CfLA-GFP. LA-GFP is labeling the actin bundle underneath the raphe. Autofluorescence of the chloroplasts is shown in red. **D:** Images of a single cell, each image is a projection of two optical sections. The actin bundles are gradually twisted around the longitudinal axis. **E:**Optical section. Additional to the actin bundle under the raphe, a finer actin network is spanning around the nucleus and further actin filaments are in the cytoplasm. N — nucleus; F — F-actin; R — actin bundle underneath the raphe. Scale bars 10 μm.

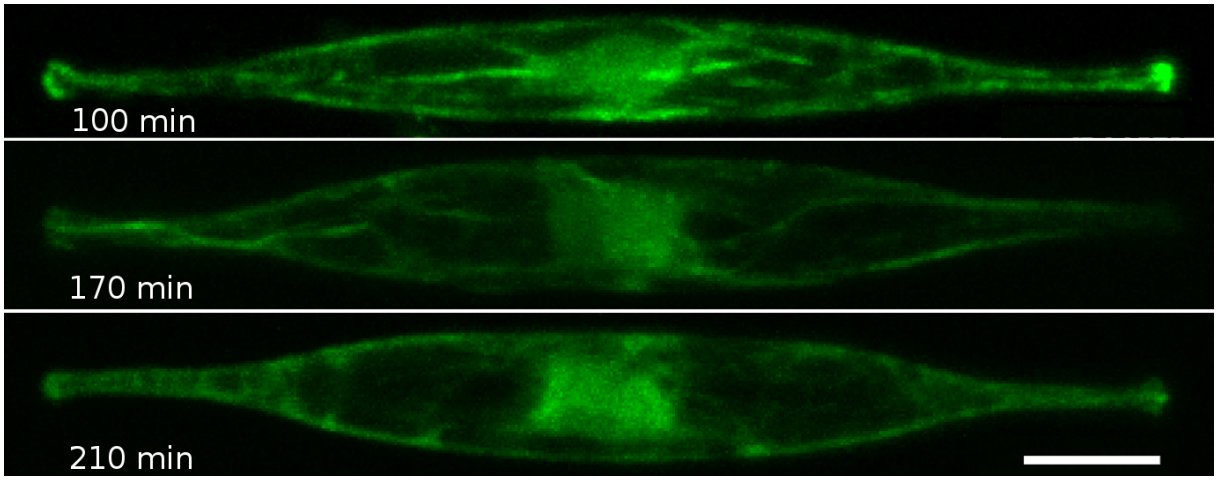


Figure 11.13: F-actin-depolymerization in *C. fusiformis* by latrunculin A. The LA-GFP reporter was visualizing the depolymerization of F-actin through latrunculin A (5 μM). Full cell projection of all optical sections, of three different cells. After 100 min the thick actin bundles started to disappear, while filaments are still visible. A treatment of 170 min depolymerized the actin bundles even stronger. 210 min treatment of 5 μM latrunculin A is effective to depolymerize actin filaments completely. Scale bar 5 μm.

12 Discussion

12.1 Transformation of *C. cuspidata*

To ensure that the protocol for diatom transformation was successfully introduced in the laboratory, *P. tricornutum* was always transformed along with the other test species as a positive transformation control (Apt et al., 1996).

C. cuspidata could not successfully be transformed with plasmids obtained from the groups of Nicole Poulsen and Peter Kroth (see Tab. 10.1), although the plasmids have previously been used in these groups to transform *P. tricornutum*, *T. pseudonana* and *C. fusiformis*. Furthermore the transformation protocol for diatoms was well established in the laboratory as transformation of *P. tricornutum* as well as *C. fusiformis* succeeded. There are various possibilities, why the approach to transform *C. cuspidata* failed. First of all *C. cuspidata* has in comparison to *P. tricornutum* and *C. fusiformis* a strongly silicified frustule. Therefore, it might be that the tungsten particles carrying the DNA, did not penetrate the thick frustule of *C. cuspidata*. This possibility, however, can be excluded as it was ascertained, that the impact of the tungsten particles is strong enough to penetrate the frustule. Cells hit by multiple tungsten particles showed perforations and became leaky. Another possibility could be, that the recovery of the cells after transformation and before plating onto the antibiotic selection plates was not sufficient. To narrow down this experimental problem, recovery conditions were modified. Cells were recovered on plates with and without liquid medium added, as well as in liquid medium for periods between one to three days. However, none of the recovery attempts were leading to success. One could furthermore argue, that the culture may have not been axenic and that bacteria overgrew the diatoms after transformation, thus making it impossible for transformants to build colonies. This, however, is also unlikely, because one of the two chosen antibiotics, G418 was lethal for all bacteria (chapter 3.3) and still did not lead to transformants. Keeping all those experimental approaches in mind, it leads to the conclusion that none of the promoters used to drive expression in *C. cuspidata* was efficient and therefore cells did not acquire antibiotic resistance. This is in accordance with previous diatom transformations, which showed that established transformation protocols work best with endogenous promoters (Apt et al., 1996; Dunahay et al., 1995; Fischer et al., 1999; Poulsen et al., 2006; Muto et al., 2013), resulting in a higher expression efficiency than heterologous promoters (Muto et al., 2013). Therefore a new plasmid for *C. cuspidata* was constructed, with an endogenous promoter driving the expression cassette.

12.1.1 Amplifying an endogenous promoter/terminator construct from a *C. cuspidata* gene

Promoter/terminator sequences are not conserved and can therefore not be directly PCR-amplified out of an unsequenced genome. Because of that, a well conserved gene has to be amplify first, with a set of degenerated primers. The second step is than to amplify the upstream and downstream regions of the conserved gene to obtain the wanted promoter/terminator sequences. When choosing

an endogenous promoter for transformation one has to take into account, that the conserved gene controlled by this promoter is ubiquitous and highly expressed throughout the cell cycle. All of these prerequisites were taken into account, choosing ubiquitin as a gene to retrieve a good promoter for the purpose to construct an expression cassette for *C. cuspidata*. Ubiquitin promoters have previously been, successfully used in other organisms for transformation (Vidali et al., 2009).

Ubiquitin promoter/terminator

The ubiquitin protein is synthesized as a precursor polyprotein consisting of either polyubiquitin chains or a single ubiquitin fused to an unrelated protein, cleaved rapidly after translation (Özkaynak et al., 1984; Lund et al., 1985). In the present case, the ubiquitin-40S ribosomal S27a gene was chosen to amplify ubiquitin out of the *C. cuspidata* genome. This pre-protein is post-translationally processed, generating free copies of a ubiquitin monomer and a ribosomal protein S27a (Tarawneh et al., 1994).

A highly degenerated forward primer (768 times degenerated), designed to match the 5'-ubiquitin coding region and two reverse degenerated primers (up to 108 times degenerated) for the S27a region, successfully amplified two ubiquitin sequences out of the genomic *C. cuspidata* DNA. The obtained sequenced matched perfectly with the ubiquitin sequences retrieve from the *P. tricornutum* genomic database.

The promoter and terminator sequences of a ubiquitin gene were successfully amplified by the genome walking method (Siebert and Chenchik, 1995; Siebert et al., 1995). A 720 bp long sequence downstream of the ubiquitin was amplified, very probably containing the full terminator sequence. Most diatom terminators of established expression cassettes are between 300 bp - 550 bp long (Zaslavskaia and Lippmeier, 2000; Apt et al., 1996), therefore the 720 bp long amplified downstream region of the ubiquitin gene should be efficient as a terminator. For the promoter sequence, only a 379 bp long fragment could be amplified. Promoters can vary in size, while most promoters are between 100 bp - 1000 bp (Karni and Felder, 2007), where for example the median size of a promoter in *Saccharomyces cerevisiae* is 455 bp (Kristiansson et al., 2009). Diatoms have been previously successfully transformed with a promoter size of 430 bp (Zaslavskaia and Lippmeier, 2000), therefore the amplified upstream sequence of the ubiquitin sequence may just be long enough for a promoter.

12.1.2 Construction of the expression cassette with an endogenous promoter/terminator

For cloning the expression cassette, the 379 bp long fragment of the putative promoter plus the 160 bp long, proximal intron, starting only 3 nucleotides after the start codon, was chosen. It was shown by Rose (2002) that introns, especially promoter-proximal introns (Rose et al., 2008) influence the expression efficiency of the promoter. Since it is known that a gap of at least 50 bp between a promoter and the reporter gene increases the expression probability of the reporter gene, a second promoter was cloned, containing the 379 bp long promoter fragment, but only 84 bp of the ubiquitin gene, containing half of the intron. For the terminator a 489 bp long sequence was cloned.

According to the determined antibiotic selection requirements (chapter 3.3), two antibiotic resistance genes were chosen as selection markers. Neo, a G418 resistance gene and natII for

ClonNat resistance. Furthermore eGFP was chosen as the reporter gene, because it has previously been shown that diatoms are only able to express eGFP and not other variants of GFPs (Zaslavskaia and Lippmeier, 2000).

12.1.3 Attempts of transforming *C. cuspidata* with the endogenous promoter/terminator construct

The transformation approach for *C. cuspidata* using the endogenous promoter to drive the expression of the antibiotic resistance gene, did not result in any transformant colonies.

For a transformed cell to survive the antibiotic screening, the resistance gene has to be strongly expressed throughout the cell cycle. It might be, that the expression rate of the resistance gene was not efficient enough, although the EST databases showed a high ubiquitin expression in *T. pseudonana* as well as in *P. tricornutum*. Since the genome of *C. cuspidata* encodes at least three different ubiquitin genes fused to S27a, it cannot be excluded that the one tested promoter here does not continuously drive the expression of ubiquitin, or it may even be a pseudogene. To overcome the possibility that no transformants were obtained due to the selection method, another approach was chosen. Therefore *C. cuspidata* transformants were visually screened for eGFP expression under the microscope with a 200-fold magnification. This experimental screening approach did not prove to be feasible. It is impossible to screen a total of $4x10^7$ *C. cuspidata* cells by eye and not overlook the potential 1 - 10 cells expressing eGFP. After all the published transformation rate for *P. tricornutum* using the same protocol leads to 10 - 100 transformants per $1x10^8$ cells (Apt et al., 1996). Because *C. cuspidata* is about 10 times larger than *P. tricornutum*, 10 times less cells can be plated within the shooting diameter, leading to a reduction of the transformation efficiency by one order of magnitude. *C. cuspidata* is not only longer than all other transformable diatoms, it also has a much greater diameter. The volume ratio between protoplast and nucleus is therefore much larger in *C. cuspidata* than in the other diatoms that have been transformed so far. This leads to the effect that the frequency of DNA-coated tungsten particles hitting the nucleus and not the cytoplasm is much lower, inevitably further reducing the already low transformation efficiency by microprojectile bombardment. In addition *C. cuspidata* has two prominent large vacuoles, together nearly spanning 2/3 of the cell. This might have also complicated the transformation process, as cells might have not been able to repair the damage, when the vacuoles were hit too often by tungsten particles. Of course it is also possible, that the promoter sequence was not long enough with only 379 bp, i.e. further important regulatory sequences may lay upstream of the amplified promoter fragment.

Up to now, it remains unsolved for what reasons the transformation of *C. cuspidata* failed. In future studies it would be necessary to choose a more highly expressing endogenous promoter, i.e. the promoter of a fucoxanthin chlorophyll binding protein, which has been successfully used before for other diatom species (Apt et al., 1996; Poulsen et al., 2006). In that context it would certainly help, if the genome of *C. cuspidata* would be sequenced.

12.2 Transformation and visualization of the actin cytoskeleton in *C. fusiformis*

C. fusiformis has been previously transformed with eGFP, although its expression was driven by a different promoter. Images off the transformants (Fig. 11.12) were available (Poulsen and Kröger,

2005) to compare the expression pattern of eGFP with LA-GFP. Therefore the two transformation approaches with CfGFP, not leading to transformants, were not repeated.

The pattern of the fluorescence signal clearly differs between eGFP and LA-GFP. eGFP is diffusely spread throughout the cell, whereas LA-GFP is labeling filamentous structures and fine meshworks, as well as thick bundles underneath the raphe. These thick bundles are a characteristic actin architecture of raphid pennate diatoms (Drum, 1963; Edgar and Pickett-Heaps, 1984). LA-GFP is clearly labeling one raphe actin bundle, instead of two raphe bundles that were recognized by confocal imaging in *C. cuspidata* (chapter 7). The cells of *C. fusiformis* are narrow with a diameter of 5 μm and more tightly packed compared to *C. cuspidata*. Regular fibulae, a characteristic feature of *C. fusiformis*, spanning over the raphe reduce the area around the raphe even further (Fig. 11.12 A). It is therefore very possible that *C. fusiformis* has two raphe actin bundles so close to each other, that they can not be separated by fluorescent light microscopy. To investigate, if two separated raphe actin bundles are present, ultrastructure studies are necessary. The latrunculin A treatment demonstrated that the fluorescence signal changed over time reducing the filamentous structures within 100 min and leading to a compleat depoymerisation of the actin cytoskeleton within 210 min. Latrunculin A binds actin monomers at a 1:1 stoichiometry and prevents them from polymerizing, thereby it reduces the free G-actin pool and consequently leads to F-actin-depolymerizing (Coué et al., 1987; Yarmola et al., 2000).

The LA-GFP fusion construct provides for the first time the opportunity to study the diatom actin cytoskeleton *in vivo*. Now the tool is available to answer questions concerning the dynamics of the actin cytoskeleton and the connection of the actin cytoskeleton to the locomotion process of pennate diatoms.

Although *C. fusiformis* is highly motile, this diatom is not an ideal model organism to study the locomotion process, because the raphe is not straight but twisted around the longitudinal axis. Here one could say "function follows form" as the structure of the raphe is leading to a continuous rotating around the longitudinal axis while the cell is gliding forward. Due to this *rotating-gliding* the organization of the cytoskeleton within the cell is very difficult to observe and especially to follow over time. Nevertheless *C. fusiformis* is currently the only available pennate diatom that is able to glide, is just large enough for cell biological observations and with established transformation techniques.

Due to time constraints, the transformation of *C. fusiformis* with the actin-specific reporter LA-GFP could not be explored any further in this thesis. However, the stable transformants build a basis for a lot of future experiments.

The data resulting from the latrunculin A treatments of *C. fusiformis* reported here, are not exhaustive, but clearly promising for a future, detailed study on actin dynamics in pennate diatoms. Some straight forward projects would be to observe the actin organization during cell division, the reaction of the actin cytoskeleton towards different inhibitors and the speed of recovery as well as the re-organization after the treatment. It would be also very interesting to record the recovery time and organization of the raphe actin bundles after inhibitor treatments, to verify the results of the fast recovery of FITC-phalloidin labeled actin bundles within under 10 sec from Poulsen et al. (1999). This would solve questions concerning the structure, polarity and dynamics of the raphe actin bundles (discussed in chapter 8).

Insights into the regulation of the actin cytoskeleton through comparative genomics

13 Introduction

13.1 Actin, Actin-related proteins & Actin-binding proteins in diatoms

Cell biologists have studies the diatoms with a focus on the cytoskeleton for more than a century (Lauterborn, 1896; Reimann, 1964; Li and Volcani, 1985; Pickett-Heaps et al., 1990), however, very little is known about regulation of the cytoskeleton and cytoskeletal interacting proteins. There is currently no publication on the presence or absence of cytoskeletal proteins except for actin, tubulin and myosin (Edgar and Pickett-Heaps, 1984; Pickett-Heaps et al., 1990; Heintzelman and Enriquez, 2010). Even with structural informations available about the diatom actin cytoskeleton (chapter 5), it is unclear how the actin network is regulated and which actin-related, actin-binding and actin-interacting proteins are encoded within the diatom genomes.

Complete genome sequences have become publicly available for four diatom species, one representative of the centric diatoms, *Thalassiosira pseudonana* (Armbrust et al., 2004) and three pennate diatoms *Phaeodactylum tricornutum* (Bowler et al., 2008), *Fragilariopsis cylindrus* and *Pseudo-nitzschia multiseries* at the Joint Genome Institute (JGI). Also the complete proteome of *Thalassiosira oceanica* (Universal protein database, UniProtKB) and expressed sequence tags (ESTs) for *P. tricornutum* and *T. pseudonana* are available (www.biologie.ens.fr/diatomics). Given the fact that other Chromalveolata genomes have meanwhile also been sequenced such as *Toxoplasma gondii* (http://toxodb. org/toxo/) and even more genomes of Stramenopiles are public available: *Aurantiochytrium limacinum*, *Schizochytrium aggregatum*, *Aplanochytrium kerguelense*, *Phytophthora sojae*, *Aureococcus anophagefferens*, at JGI as well as *Ectocarpus siliculosus* (Sterck et al., 2012), a large resource of genome information is providing excellent opportunities to perform *in silico* comparative genome analyses. Retrieving information from these resources, this study presents results of a whole genome analysis of actin, actin-related proteins and actin-binding proteins in the diatoms. Myosins have not been included in this study, as they have been already described for *P. tricornutum* by Heintzelman and Enriquez (2010).

In the following, the state of knowledge of the investigated proteins and protein families is briefly reviewed, before the results are presented.

13.2 Actin & Actin-related proteins

Actin is one of the most highly conserved and abundant proteins in eukaryotic cells and a major component of the cytoskeleton. Dynamic assembly and disassembly of the actin cytoskeleton is essential for many cellular functions such as cell motility, the control of cell shape and polarity, cytokinesis and intracellular transport as well as endo- and exocytosis. Actin is a self-assembling protein that undergoes polymerization and depolymerization under physiological conditions. Throughout the

kingdom of eukaryotes, the copy number of actin genes in the genomes is quite divergent ranging from a single copy in fungi, algae, and some protists (McDowell et al., 1996), to 10 copies in protists and metazoans including animals (Bhattacharya et al., 2000), up to a dozen and more copies in plants (Šlajcherová et al., 2012). There are even some higher plants with up to 100 actin genes, such as the ornamental species Petunia (Baird and Meagher, 1987).

The globular (G-) actin consists of four sub domains. The nucleotide and a divalent cation (Mg^{2+} or Ca^{2+}) are bound between subdomains 2 and 4 (minus-end) (Fig. 13.1) within the following residues: Asp^{11}, Ser^{14}, Gly^{15}, Lys^{18}, Asp^{115}, Asp^{158}, Gly^{159}, Val^{160}, Lys^{214}, Glu^{215}, Gly^{304}, Thr^{305}, Met^{307}, and Lys^{339} (Kabsch et al., 1990). This cleft has an enzyme activity capable of hydrolyzing ATP. The opposite end of the molecule (plus-end) is built up by subdomains 1 and 3, forming a hydrophobic cleft between them.

G-actin, to which ATP is bound, has the ability to polymerize to filamentous (F-)actin. The filaments are helically constructed with a fast growing (plus) end and a slow growing (minus) end. The rate of polymerization at the plus-end is higher than the rate of depolymerization and at the minus-end it is the other way around. After polymerization ATP is slowly hydrolyzed to ADP+Pi and the inorganic phosphate is released (Fig. 13.2) converting the F-actin into the ADP-bound stage, which is less stable than the ATP-bound stage and eventually depolymerizes.

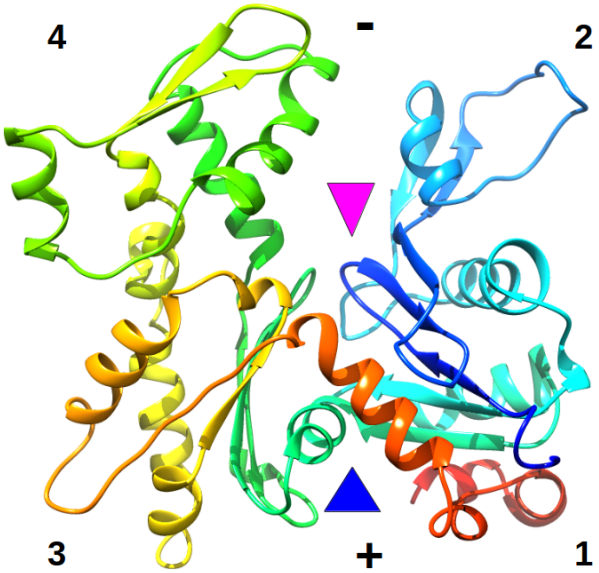


Figure 13.1: Predicted tertiary structure of G-actin using the diatom actin (Phaeo 51157) as an example. The plus-end (+) is at subdomains 1 and 3 and the minus-end (-) at subdomains 2 and 4. The nucleotide and divalent cation binding site within the upper cleft is indicated with a pink arrowhead. The hydrophobic cleft is the major interaction site for ABPs (blue arrowhead) at subdomains 1 and 3. Tertiary structure is drawn in rainbow colors from blue (N-terminus) to red (C-terminus).

However, the rate of filament turnover is much slower *in vitro* than *in vivo*. The structure and dynamics of the actin cytoskeleton *in vivo* are largely dependent on, and tightly regulated by actin-binding proteins (ABPs) which assist in nucleation, polymerization/depolymerization, fragmentation as well as branching of the actin filament (Pollard et al., 2000; Pollard and Borisy, 2003; Cooper and Schafer, 2000; Condeelis, 2001; Blanchoin et al., 2010, 2014). Other actin-associated proteins are important for stabilization, bundling and interlinking of actin filaments (dos Remedios et al., 2003; Thomas et al., 2009). In most cases the action of ABPs is dose dependent and many of them compete for the same binding site on the actin molecule. Most prominent in this respect is the hydrophobic cleft between subdomains 1 and 3 (Dominguez, 2004), which is also the contact site between adjacent actin molecules in the filament (Fujii et al., 2010; Oda et al., 2009). Most ABPs have an α-helix domain that interacts with conserved residues within the hydrophobic cleft of actin (i.e., Tyr^{143}, Ala^{144}, Gly^{146}, Thr^{148}, Gly^{168}, Ile^{341}, Ile^{345}, Leu^{346}, Leu^{349}, Thr^{351}, Met^{355}, Fig. 13.1) and probably the C-terminus of actin (Dominguez, 2004).

Among the proteins that interact with actin are the so-called actin-related proteins (ARPs). The members in this family are structurally very similar to actin, numbered according to their decreasing homology with actin (Arp1 to Arp11); Arp1 is structurally most closely related to actin and Arp11 the most divergent from actin (Poch and Winsor, 1997). The ARPs can be grouped

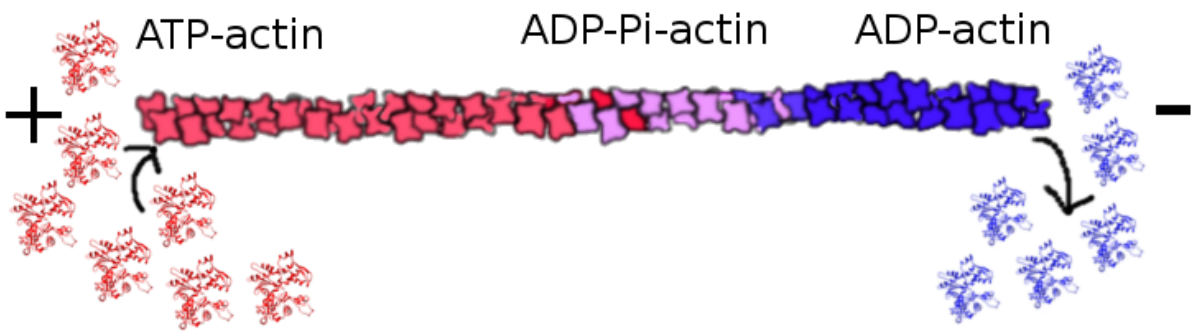


Figure 13.2: Drawn illustration of the actin filament polymerization and depolymerization. ATP bound G-actin (red) polymerizes into F-actin with a helical twist by binding to the plus-end of an existing filament. G-actin hydrolysis ATP to ADP-Pi (purple) leading to a less stable filament. In the ADP-bound state (blue) the filament eventually depolymerizes. The polymerization at the plus-end while depolymerization at the minus-end produces a phenomenon known as treadmilling.

into cytoplasmic ARPs (Arp1-3 and Arp10-11) and nuclear ARPs (Arp4-9). Regardless of their cellular location, ARPs are often found as homopolymers or heteropolymers and assemble with other proteins into large multiprotein complexes that frequently associate with actin (Schafer and Schroer, 1999). Nuclear ARPs are important in the context of gene transcription as well as chromosomal organization (Schafer and Schroer, 1999; Kast and Dominguez, 2011). The cytoplasmic Arps have different functions, i.e. the Arp2/3 complex, together with half a dozen other proteins, has been known for more than a decade as the molecular key player in actin filament branching and plus-end F-actin nucleation (Mullins et al., 1998; Beltzner and Pollard, 2008; Volkmann et al., 2014). The other cytoplasmic multiprotein machinery involving ARPs is the Arp1- and Arp11-containing dynactin complex, which activates and targets a microtubule dependent dynein activity (Schafer and Schroer, 1999).

13.3 Actin-binding proteins

It is estimated that eukaryotic cells have over 100 different ABPs to regulate the dynamics and network properties of the actin cytoskeleton (dos Remedios et al., 2003). ABPs can regulate the actin cytoskeleton in various ways by binding actin monomers and thereby influencing the G-actin pool; by binding to the filaments and thereby altering the stability of F-actin; by capping either the plus or the minus-end of the filament and thereby influencing the assembly/disassembly kinetics; by creating breaks or even severing the filaments and thereby regulating the network properties. The plethora of the ABPs can be clustered into different classes, indicating their likely mechanism of action.

13.3.1 Monomer binders & Depolymerizing proteins

Continuous subunit turnover is essential for the dynamics and plasticity of the actin cytoskeleton in order to allow rapid remodeling in response to internal and external signals. Cellular processes such as cell motility, control of cell shape and polarity as well as cytokinesis are driven by finely tuned actin polymerization, which requires the availability of a suitable sized ATP-actin pool. In oder to maintain a large pool of ATP-actin, aged filaments from the existing filament network need to be broken down and depolymerized. The resulting ADP-actin monomers then have to be recycled

back to an assembly-competent ATP-bound state. The major rate-limiting steps in actin turnover, filament disassembly, and monomer recycling are accomplished by the combined action of a small set of proteins: ADF/cofilin, Aip1, coronin, Cyclase-associated proteins (CAP), profilin and a few others. These proteins are highly abundant and well conserved in most eukaryotes examined so far, showing their fundamental importance for actin cytoskeletal regulation. The only exception is coronin, which is absent from plants and protists (Morgan and Fernandez, 2008).

ADF/cofilin binds ADP-subunits in F-actin and thereby severs filaments (Bamburg, 1999; Carlier et al., 1999; Maciver and Hussey, 2002; DesMarais et al., 2005; Van Troys et al., 2008). Further more ADF/cofilin bound to F-actin enhance Pi release from neighboring subunits, thereby promotes the aging of the filament leading to depolymerization (Blanchoin and Pollard, 1999). Aip1 alone has a negligible effect on actin filament dynamics, but it is a unique regulator of ADF/cofilin. Interacting with ADF/cofilin, Aip1 enhances the disassembly of actin filaments and therefore increases the ADF/cofilin mediated filament turnover (Rodal et al., 1999; Ono, 2003; Voegtli et al., 2003; Mohri et al., 2004; Kueh et al., 2008). An interaction between the three proteins Aip1, cofilin and coronin causes filament disassembly in abrupt bursts from both ends (Kueh et al., 2008). The mechanism of cofilin-coronin interaction is incompletely understood and may differ between species.

The enhanced F-actin disassembly through Aip1, cofilin and coronin leads to the rapid accumulation of ADP-actin monomers bound to ADF/cofilin. Profilin is well known to accelerate the rate of ADP to ATP exchange in G-actin (Goldschmidt-Clermont et al., 1991; Mockrin and Korn, 1980; Korenbaum et al., 1998; Yarmola and Bubb, 2006) and binds to G-actin at a molar ratio of 1:1 (Schutt et al., 1993; Vinson and Archer, 1993). However, profilin can only weakly promote nucleotide exchange of ADF/cofilin-bound ADP-actin monomers (Balcer et al., 2003) and totally lacks the nucleotide exchange activity in plants and green algae (Kovar et al., 2000, 2001). These observations raise the question whether profilin has a major role in recharging ADF/cofilin-bound ADP-actin monomers, or if other proteins i.e. CAPs are overtaking that role. CAPs are also regulating the Ras–cAMP signaling pathway, but the present study only emphasizes on their roles in regulating the actin cytoskeleton. CAPs are G-actin sequestering proteins that inhibit spontaneous actin polymerization, as well as accelerate the nucleotide exchange of ADP-actin to ATP-actin even in plants (Chaudhry et al., 2007; Hubberstey and Mottillo, 2002; Ono, 2013). In contrast to profilin which binds ATP-actin with a higher affinity (Pantaloni and Carlier, 1993), CAPs bind ADP-actin with a relatively high affinity and therefore directly compete with ADF/cofilin for the ADP-actin pool (Mattila and Quintero-Monzon, 2004; Moriyama and Yahara, 2002).

13.3.2 Anchors to membranes

The plasma membrane is shaped by the underlying protein meshwork of the actin cytoskeleton, the so called cortical cytoskeleton. The cortical cytoskeleton is highly dynamic and capable of rearranging in response to different stimuli. ABPs play a crucial role in anchoring the actin cytoskeleton to the cell boundary, mediating the interaction between actin filaments and the plasma membrane.

Annexin

Annexins are a super-family of ubiquitous, calcium- and membrane-binding proteins within a wide range of eukaryotes. Annexins can be divided into five families defined by their distribution in

the major phylogenetic branches of the eukaryotic kingdom: A (vertebrates), B (invertebrates), C (fungi and chromalveolates), D (plants) and E (protists) (Moss and Morgan, 2004). It is generally estimated, that annexins have a fundamental biological role, involved in diverse membrane related processes (Blackbourn and Battey, 1993; Carroll et al., 1998; Konopka-Postupolska, 2007; Konopka-Postupolska et al., 2011). Several vertebrate annexins bind to F-actin in a Ca^{2+} dependent manner (Tzima et al., 2000) and some annexins even induce actin bundling (Filipenko and Waisman, 2001). Annexins are not only acceptable to bind to F-actin, but also interact with membranes and are therefore assuming a linking role in the cell between the cytoskeleton and the membrane dynamics, as well as in membrane trafficking.

13.3.3 Capping proteins

The dynamics of the actin filament assembly and disassembly is also regulated by proteins binding to the plus- or minus-end of the filament. A group of proteins, i.e. gelsolin, severin, profilin, villin, CapG, tensin, fragmin and CAPZ bind to the plus-ends of the actin filaments (Schafer et al., 1996; dos Remedios et al., 2003). These capping proteins block the exchange of subunits and thereby terminate the elongation at the plus-end of the filament (Schafer et al., 1996), leading to a regulation of the assembly at the filament and efficiently nucleate F-actin growth at the minus-end (Cooper and Pollard, 1985; Xu et al., 1999; Blanchoin et al., 2000; Huang et al., 2003; Galland et al., 2013). Capping proteins not only posses capping but also severing activities, the only exception is CAPZ, which is not effecting the rate of fragmentation (Casella and Torres, 1994). Apart from the capping activity, CAPZ also stabilizes small actin oligomers leading the actin nucleation (Cooper and Pollard, 1985; Caldwell et al., 1989).

Formin

Ubiquitous among eukaryotes, formins are a large family of multidomain proteins defined by the possession of the conserved formin homology2 (FH2) domain. Formins are multi-functional proteins involved in various aspects of cytoskeletal organization. They are required for actin cable assembly (Sagot et al., 2002a), able to nucleate actin filaments (Evangelista et al., 2002; Pruyne et al., 2002; Sagot et al., 2002b,a), act as *leaky* plus-end caps (Xu et al., 2004; Otomo et al., 2005), anchor filaments to membranes (reviewed in Cvrčková, 2013) and cooperate with profilin to promote rapid filament elongation (reviewed in Paul and Pollard, 2009). Formins are cytoplasmic proteins, but they are also transported into the nucleus, where they influence the nuclear-actin dynamics (Baarlink and Grosse, 2014).

The signature FH2 domains of two formins form a unique "tethered dimer", each FH2 domain of the dimer is competent to interact with an actin subunit at the plus-end of a filament (Xu et al., 2004; Otomo et al., 2005). In most formins a proline-rich region, the formin homology 1 (FH1) domain, is located N-terminally of the FH2 domain. The FH1 domain is described to be essential for the physiological functions of fomins, being able to bind and interact with profilin-actin (Chang et al., 1997; Watanabe et al., 1997; Courtemanche and Pollard, 2012). Formin-profilin-actin interaction at the plus-end of the filament remarkably accelerates the elongation rate beyond that of free actin filaments (Sagot et al., 2002b; Romero et al., 2004; Paul and Pollard, 2008). Less is known about the function of the two described formins, *D. discoideum* formin ForC and the malaria parasite *Plasmodium* formin MISFIT, that lack the FH1 domains (Kitayama, 2003; Higgs and Peterson, 2005; Rivero et al., 2005; Bushell et al., 2009).

13.3.4 Bundlers & Crosslinkers

Actin-bundling proteins like villins, fascin, fimbrin, filamin, formin, alpha-actinin and scruin are necessary to generate tightly packet bundles of F-actin in a cell. Fascins are a structurally unique group of actin cross-linking proteins consisting of four β- trefoil domains capable to bundle thick filaments i.e. in filopodia (reviewed in Jayo and Parsons, 2010).

Gelsolin Superfamily Proteins

Within the Gelsolin superfamily of actin-binding proteins divers functions have been observed such as capping/uncapping of F-actin, nucleating and bundling of F-actin, as well as severing of F-actin in a Ca^{2+} dependent manner (reviewed in McGough et al., 2003). All members of the family have conserved gelsolin repeat domains (Way et al., 1990; McLaughlin et al., 1993). The smallest members (severin, fragmin and CapG) contain three of these conserved gelsolin repeats (S1–S3), whereas most others have six repeats: gelsolin, villin, adseverin, advillin, protovillin and flightless (reviewed in Puius et al., 1998). Although the small proteins severin, fragmin and CapG have only three gelsolin domains they are fully capable of the severing, capping and nucleating activities just like the proteins with the six repeats (André et al., 1988; Yin et al., 1990; Eichinger et al., 1991). The only exception is CapG, which is not showing any severing activities (Zhang et al., 2006). Some family members i.e. CapG (Renz and Langowski, 2008) are localized in the nucleus where they adopted specific functions, including the regulation of transcription.

In addition to the gelsolin repeat some protein of the superfamily contain a villin headpiece at the C-terminus (VH), which correlates with the crosslinking and bundling activities (Glenney et al., 1981; Pope et al., 1994) (reviewed in Khurana and George, 2008). Dematin a membrane-cytoskeleton-associated protein with F-actin-binding activity contains a homologous domain to the VH at the C-terminus (Rana et al., 1993) that induces F-actin bundle formation and stabilization.

14 Material & Methods

Organism specialized sequence databases of the Joint Genome Institute (www.jgi.doe. gov/), the Blastocystis Genome Browser (www.genoscope.cns.fr/externe/GenomeBrowser /Blastocystis/), the Toxoplasma Genomics Resource (http://toxodb.org/toxo-release4-0/home.jsp) and the Online Resource for Community Annotation of Eukaryotes (http://bioinformatics.psb.ugent.be/ orcae/overview/Ectsi) were accessed for comprehensive ge-nome and protein BLAST searches for following Chromalveolata: *Thalassiosira pseudonana*, *Phaeodactylum tricornutum*, *Fragilariopsis cylindrus*, *Pseudo-nitzschia multiseries*, *Aurantiochytrium limacinum*, *Schizochytrium aggregatum*, *Aplanochytrium kerguelense*, *Phytophthora sojae*, *Phytophthora infestans*, *Aureococcus anophagefferens*, *Ectocarpus siliculosus* and *Toxoplasma gondii*. The proteins sequences of *Thalassiosira oceanica*, *Mus musculus*, *Dictyostelium discoideum* and *Arabidopsis thaliana* were retrieved through NCBI and HMMER search tools (hmmsearch)(http://hmmer.janelia.org/).

Up to 15 protein sequence representing major taxa including mammals, fish, plants, flies, amphibia, worms, fungi, and Stramenopiles were used to screen for actin, Arp1 to Arp11 and ABPs in the databases described above. Additional diatom sequences were retrieved using bioinformatic search tools such as BLAST (blastp,tblastn, blastn, blastx) at NCBI (nr database), and Profile hidden Markov models (HMMER) (http://hmmer.janelia. org/). PSI-BLAST and HMMER search tools (hmmsearch) were used to confirm the absence of profilin in diatoms.

Truncated protein sequences from *T. pseudonana* and *P. tricornutum* were compared with the corresponding sequences from the EST databases (Maheswari et al., 2005, 2008) (http://www.diatomics.biologie.ens.fr/EST3/). Proteins with a truncated amino acid sequences were viewed in nucleotide and 3-frame translation and C-terminal elongated until the next stop codon. N-terminal elongation was performed until the first ATG following a stop codon. In the case were the search result of amino acid sequences had due to automatic annotation a wrong intron-exon structure, nucleotide sequences were translated using the Sequence Manipulation Suite web server (www.bioinformatics.org/sms2/ index.html). All obtained sequences were blasted in the NCBI database (www.ncbi.nlm. nih.gov/), aligned and manual corrected to determine the correct exon structure of the protein sequence.

14.1 Sequence alignments

The candidate actins, ARPs and ABPs (Tab. 20.1) were compared with a larger repertoire of actins, ARPs and ABPs from a variety of eukaryotes retrieved from NCBI, UniProtKB (http://www.uniprot.org/uniprot/) and the Protein Data Bank (PDB) (http: //www.rcsb.org/ pdb/home/home.do). All sequences were aligned, depending on the demand with the Clustal Omega, MAFFT, Kalign, or MUSCLE software from EBI (www.ebi.ac.uk/Tools/msa/clustalo/) using pairwise parameters set at default settings. The aligned candidate ARP sequences were screened for the describing signatures of the ARP subfamilies. The identification and clustering of the ARP proteins into

the ARP subfamilies was further verified through the Actin Related Proteins Annotation server (ARPAnno)(Muller et al., 2005).

14.2 Phylogenetic analysis

The actin and actin-related proteins of diatoms were assembled into a multiple sequence alignment using the MUSCLE utility of the Molecular Evolutionary Genetics Analysis (MEGA) software version 5.05. The phylogenetic tree was constructed with 2000 bootstrap replicates using the Maximum Likelihood method based on the Whelan and Goldman model with a discrete Gamma distribution (+G, parameter = 8.0938)(Tamura et al., 2011; Whelan and Goldman, 2001).

14.3 Structural studies

The secondary structure was predicted through the Protein Homology/analogY Recognition Engine V 2.0 (Phyre^2) and the PSIPRED Protein Sequence Analysis Workbench (http://bioinf.cs.ucl.ac.uk/psipred/). Homology detection and structure prediction by HMM-HMM comparison (HHpred) (http://toolkit.tuebingen.mpg.de/hhpred) and Phyre^2 predicted the 3D structure of the identified proteins to a template sequence. The crystal structure templates with the following PDB-ID were used: 1YAG — actin, 4B1Y/2FXU/1K8K — Arp1; 1K8K — Arp2; 3QB0 — Arp4; 4B1Y/3DWL/2FXU/4FO0 — Arp6; 1NRO — Aip1; 1QPV — ADF/cofilin; d1hm6a — Annexin; 1k8F — CAP; 1IZN — CAPZ; 2AQ5 — Coronin. Using the software UCSF Chimera (Pettersen et al., 2004) the PDB models of the proteins were visualized and the molecular structure analyzed. For structural comparison the PDB sequences were superimpose with the MatchMaker tool of Chimera or the Dali Database (Holm L, 2010). To display the structures in tiled arrangement the Tile Structure tool of Chimera was selected.

14.4 Motifs & patterns

Scanning proteins for motifs and pattens was achieved by Motif Scan, ExPASy the SIB Bioinformatics Resource Portal (http://myhits.isb-sib.ch/cgi-bin/motif-scan), Conserved Domains from NCBI (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Marchler-Bauer and Bryant, 2004; Marchler-Bauer et al., 2009, 2011) and the SMART software (http://smart.embl-heidelberg.de/). InterPro predicted domains and important sites and additional classified the protein sequences into families. The results from the different portals were compared, evaluated and graphical illustrated with the MyDomains - Image Creator from ExPASy (http://prosite.expasy.org/cgi-bin/prosite/mydomains/).

Databases chosen to scan for motifs and patterns by Motif Scan were the following: PeroxiBase profiles [perox], HAMAP profiles [hamap], PROSITE patterns [pat], More profiles [pre], Pfam HMMs (local models) [pfam fs], Pfam HMMs (global models) [pfam ls], PROSITE patterns (frequent match producers) [freq pat] and PROSITE profiles [prf].

15 Results & Discussion

The genomes of the four diatoms *F. cylindrus*, *P. multiseries*, *P. tricornutum*, *T. pseudonana*, and the proteome of *T. oceanica* were screened for actin, 11 ARPs and 23 ABPs. The genomes of the screened diatoms vary significant in size, with two small genomes of about 30 Mb (*P. tricornutum* and *T. pseudonana*) and three larger genomes of at least 80 Mb (*F. cylindrus*, *P. multiseries* and *T. oceanica*) (Tab. 15.1). The two small genomes of *T. pseudonana* and *P. tricornutum* were the first sequenced and since than are the best studied diatom genomes, whereas the three larger genomes are not yet annotated and for *T. oceanica* only the proteome is public available.

In general the BLAST results from *T. oceanica* are to be treated with caution. The search results through NCBI often resulted in fragmented sequences and may therefore not always represent complete reading frames. As the genome of *T. oceanica* is not public available through JGI or other institutions, full length sequences could not be obtained.

15.1 Actin in diatoms

BLAST searches revealed one actin protein for *T. pseudonana*, *T. oceanica*, *F. cylindrus* as well as *P. multiseries*, and two actin proteins for *P. tricornutum*. The genome of the centric diatom *T. pseudonana* contained only a single copy of the actin gene, whereas all screened pennate diatoms had at least one duplicate in their genome. The actin gene copies within a species were all similar, but not identical. The actin gene family in *F. cylindrus* was leading the field with six duplicated actin genes, all were highly similar gene copies with an exon identity of at least 97.3% between them. All nucleotide exchanges within the exons were conservative, coding for the same protein. A similar situation was encountered in *P. multiseries*, where the two actin-coding sequences showed 98.77% identity on nucleotide level, but 100% identity on amino acid level. The genome of *P. tricornutum*

Table 15.1: Genome statistics of the five screened diatom species Data retrieved from JGI and (Lommer et al., 2012). All pennates are isogamous non-flagellated and the centrics are oogamous flagellate. Fc — *Fragilariopsis cylindrus*, Pm — *Pseudo-nitzschia multiseries*, Pt — *Phaeodactylum tricornutum*, Tp — *Thalassiosira pseudonana*, To — *Thalassiosira oceanica*, *— pleomorphic.

	Fc	Pm	Pt	Tp	To
Diatom lineages	Pennate	Pennate	Pennate*	Centric	Centric
Nuclear genome Size (Mb)	80.5	218.73 (Genome Assembly size)	27.4	32.4	81.6 estimated
Predicted genes	18,077	19,703	10.402	11.776	10.109

Table 15.2: Resulting percent identity matrix of diatom actins. Alignment on the amino acid level – created with Clustal2.1.

	Thaps Act	Phatr Act1	Phatr Act2	Thaoc Act	Psemu Act
Fracy Act	93.63	97.88	97.35	97.35	99.20
Psemu Act	93.90	98.14	97.08	97.61	
Thaps Act	95.76	98.94	98.41		
Phatr Act1	94.69	98.94			
Phatr Act2	95.23				

had one additional actin gene copy, and was furthermore the only diatom genome, that encodes two true actin isoforms (*PhatrAct1* and *PhatrAct2*). These were 84.66% identical on nucleotide level translating to 98.94% identity on the amino acid level. These minor differences relate to only four amino acid exchanges (Ser^2 to Ala^2, Asp^4 to Glu^4, Ala^{233} to Ser^{233}, Cys^{269} to $Thr2^{269}$). Of these only the last exchange was non-conservative, which could potentially have a functional impact, but neither this residue nor the others lay within protein-protein interacting domains. As judged from the homology matrix (Tab. 15.2) the actin genes from *P. tricornutum* seemed to group slightly better with the genes of the centric diatom *T. pseudonana* than with the actin genes retrieved from the other pennate diatom genomes. Overall, the observed low number of actin genes per genome fits well with the previous reports of other single celled organisms, including unicellular algae as well as fungi, which also contain just one or two copies of actin genes (McDowell et al., 1996; Šlajcherová et al., 2012).

The diatom actin genes code for proteins of constant length (377 aa), folding into a secondary structure of 13 β-strands and 21 α-helices (Fig. 15.1). Depending on the program used for predicting the secondary structure, the number of helices and β-strands slightly varied, the present result was based on a combined $Phyre^2$ and Chimera software analyses.

The minor differences between the six diatom actin proteins (Tab. 15.2) resulted in a highly conserved 3D-structure, with virtually no structural differences in the predicted tertiary structures (Fig. 15.2 A). The hydrophobic cleft located between subdomains 1 and 3 (Fig. 15.2 A, blue arrowhead) is the major interaction site for ABPs (Fujii et al., 2010; Oda et al., 2009; Dominguez, 2004). The screened residues in the hydrophobic cleft, potentially important for the interaction with ABPs were all conserved (Fig. 15.1, Fig. 15.2 A black arrowhead) and most of the residues in the upper cleft (between subdomain 2 and 4) responsible for ATP binding were also conserved (Fig. 15.1, Fig. 15.2 A, pink arrowhead).

According to Schutt et al. (1993) profilin binds to the hydrophobic cleft, mediated specifically by 10 residues located in subdomain 1 and 11 residues in subdomain 3. All except one (Gln^{354} to Ser^{354}) of these residues were conserved in the diatom actins (Fig. 15.1, Fig. 15.2 A, arrow). Interestingly, the single residue exchange removing a bulky, positively charged residue from the polypeptide chain was conserved between all diatoms. This could be sufficient to affect the function of this domain, although in a more delicate fashion, because other ABPs utilize the same region of the actin protein for interaction. Overall, it is likely that there is a high selection pressure on maintaining the functionality of this domain throughout evolution. It remains to be seen, whether

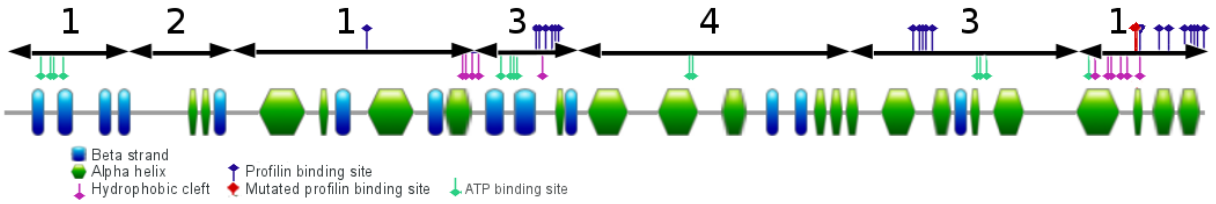


Figure 15.1: Calculated and graphical illustrated predicted secondary structure of the diatom actins (lower trace), aligned with the subdomain divisions indicated by arrows (upper trace). The positions of amino-acid residues in the hydrophobic cleft involved in ABP-binding and the ATP binding site are indicated. Note that one of the residues in subdomain 1, potentially involved in profilin binding, is mutated (red flag).

this single amino acid exchange is of significance although it is unlikely, because, as will be shown below (Fig. 15.9), profilin is completely missing from the diatoms.

Almost all plant actin genes possess an intron located shortly after the start codon (McDowell et al., 1996). This was also the case in the diatom actins, where the intron varied slightly in size (Fig. 15.3). However, two genes differed more strongly in their exon-intron structure. *ThapsAct* had two introns in the 5'-region and *PhatrAct2* lacked any intron. As already shown from the amino acid comparison *P. tricornutum* codes for two actin proteins, this result was further supported by the different intron-exon structure of these two actin genes. The intron loss in *PhatrAct2* indicates an incident of a relatively recent retrograde insertion of a processed actin mRNA transcript.

The differences in intron number and position between the centric diatom *T. pseudonana* and the three pennates (Fig. 15.3) is consistent with the evolutionary split into the two lineages. Unexpected was the over all low number of introns within the diatom actin genes, for comparison plant actin genes contain about 4 introns and metazoan actin genes contain up to seven introns (Sheterline and Sparrow, 1994; McDowell et al., 1996).

The six copies of the *F. cylindrus* actin gene seem to originate from duplication events. They were identical on the amino acid level, had the same gene model structure and the genes were located on two scaffolds with regular spacing of about 10 kbp in between. Furthermore the intron sequences in the *F. cylindrus* actin gene copies did show considerable variations. This is consistent with the well known phenomenon that introns have a faster mutation rate, because they are subjected to a much lower selection pressure than exons (Li, 1997). Interestingly, actin gene duplication is observed here only within the pennate diatoms but not within the centric species *T. pseudonana* (Tab. 15.1). This feature correlates with a striking difference in the actin cytoskeleton organization between the two diatom lineages. In addition to the actin filament network throughout the cell, pennates possess two tightly packed actin bundles running parallel to the raphe. These actin bundles are highly specialized and functionally important for cell locomotion (Edgar and Pickett-Heaps, 1984; Poulsen et al., 1999). Additional to the obvious differences in cytoskeletal organization the two diatom linages differ with respect of their morphology and sexual reproduction as pennates are isogamous and non-flagellated, whereas the centrics are oogamous with flagellated male gametes (Chepurnov and Mann, 2004; Idei et al., 2013; Sato et al., 2011). Therefore, actin gene duplication might not only be relevant for the vegetative stage, but also for the generative stage with regard to gamete morphology and behavior. However, in order to confirm the actin gene duplication in the pennates as a general trend, the genomes of further centrics and pennates will have to be analyzed. Apparently no genomic sequence informations have been available through search engines from *T. oceanica* to include in this research and therefore could not confirm the trend of *T. pseudonana*.

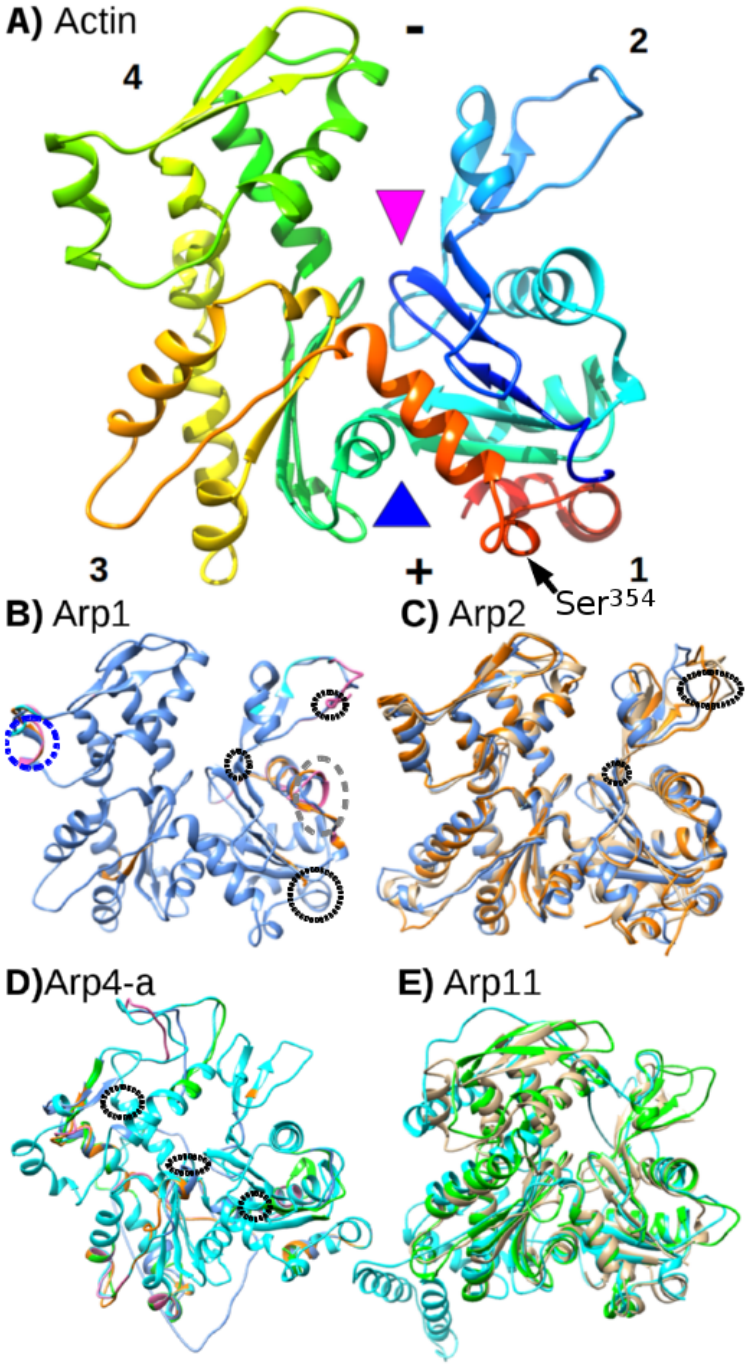


Figure 15.2: Results of the predicted tertiary structures of diatom actins and ARP subfamilies. The ARP sequences are show in the following color code: *T. pseudonana* — cyan, *P. tricornutum* — pink, *F. cylindrus* — orange, *P. multiseries* — blue, *T. oceanica* — green, and known crystal structures from other organisms are shown in gold. Deletions within a domain are indicated with blue dotted circles and insertions with black dotted circles. Regions of special interest are shown in gray dotted circles. **A)** Superimposed 3D structures of all five diatom actin proteins, no differences was observed within the tertiary structure. Subdomains 1-4 are indicated by their numbers. The nucleotide and divalent cation binding site within the upper cleft is indicated (pink arrowhead). The hydrophobic cleft is the major interaction site for ABPs (blue arrowhead). Mutated profilin-binding residue Ser^{356} is indicated in subdomain 1 (arrow). The backbone trace is colored from blue (N-terminus) to red (C-terminus). **B)** Superimposed structures of Thaps Arp1, Phatr Arp1, Fracy Arp1 and Psemu Arp1. Arp1 is structurally quite conserved with only minor differences in each subdomain. The N-terminal insertion was not structurally predicted by the (Phyre^2) software, gray circle. **C)** Superimposed structures of *Mus musculus* Arp2 4JD2-B in gold, with Fracy Arp2 and Psemu Arp2. **D)** Superimposed structures of Thaps Arp4-a, Phatr Arp4-a, Psemu Arp4-a, Fracy Arp4-a1, Fracy Arp4-a2 and Thaoc Arp4-a. **E)** Crystal structure of *Oryctolagus cuniculus* actin (3HBT) in gold, superimposed with *T. pseudonana* Arp11-like (blue) and *T. oceanica* Arp11-like (green).

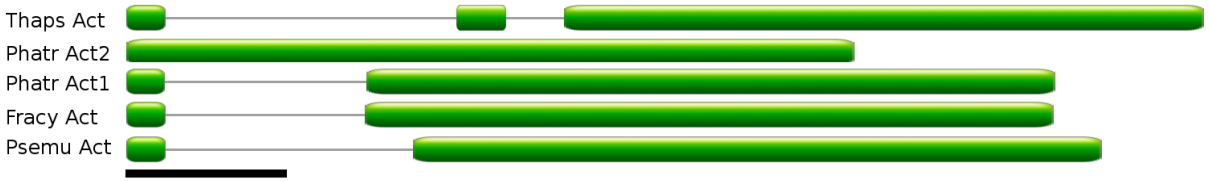


Figure 15.3: Schematic presentation of the intron-exon structure of the diatom actin gene models. Green — exon; gray — intron; Scale bar — 100 bp.

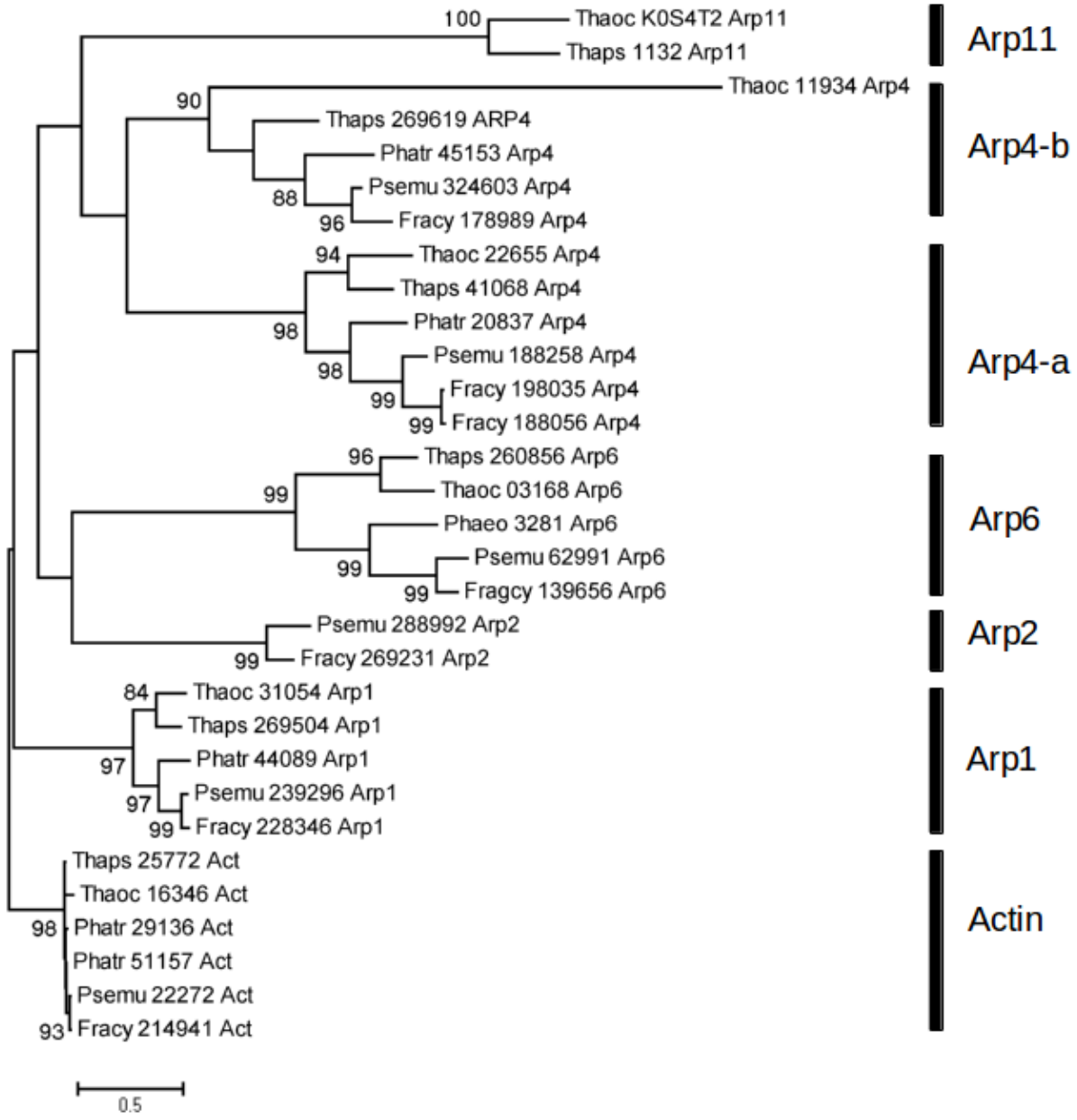


Figure 15.4: Phylogenetic tree of actin and actin-related proteins of diatoms. Bootstrap values are indicated at the branching points. For a phylogenetic analysis of *P. tricornutum* actin in the larger context with other Stramenopiles and protist phyla see (Tanifuji and Archibald, 2010).

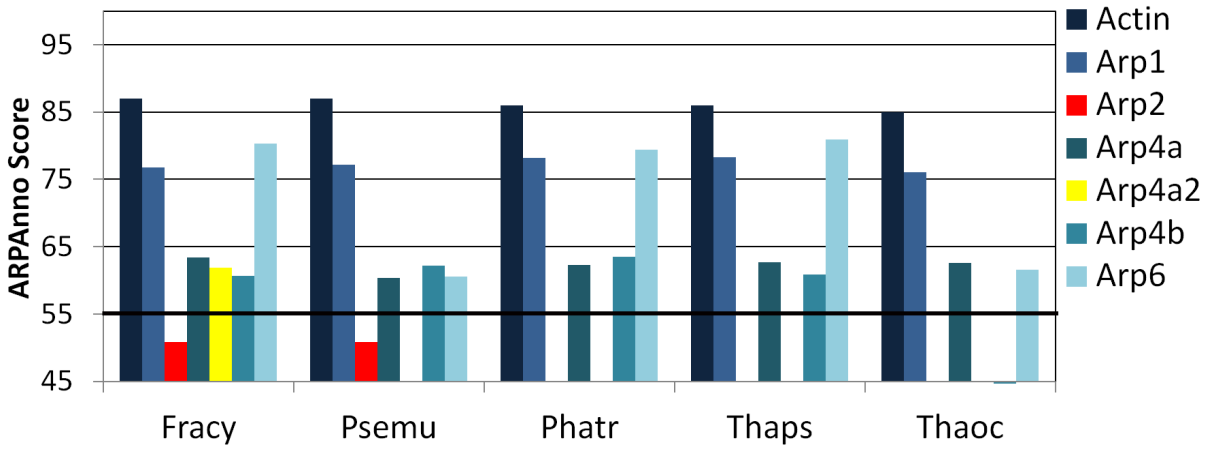


Figure 15.5: Diatom homologues of actin and ARPs. Results of the annotation score retrieved through the ARPAnno web server (Muller et al., 2005) the score of the proteins that group with an ARP subfamily is given by a confidence value on a scale from 0 to 100. A score greater than 55 (black line) is considered highly reliable. All proteins are above this threshold, with the exception of the Arp2-like proteins found in *F. cylindrus* and *P. multiseries* . Note that Arp4a is represented by two closely related isoforms in *F. cylindrus*, whereas all other species possess only one isoform. The Thaoc Arp4-b protein is C-terminally truncated and therefor not listed.

15.2 Actin-related proteins in diatoms

Screening the four diatom genomes for ARPs resulted in the identification of genes homologous to Arp1, Arp2, Arp4, Arp6 and Arp11. A single representative of the Arp1 subfamily was found in each of the diatom proteome/genomes on the basis of Arp1 specific sequence signature as confirmed by the Actin Related Proteins Annotation server (ARPAnno) (Fig. 15.5). All five Arp1 representatives were conserved between the diatom species. As a special feature, diatom Arp1s carried four additional short insertions and one deletion, all similar in size and position, which sets them apart from the non-diatom Arp1 members (Muller et al., 2005). Three of the insertions occurred in subdomain 1 and one insertion was located in subdomain 2 (Fig. 15.6, Fig. 15.2 B). However, there were also some differences between the diatom Arp1 sequences. I.e., *T. pseudonana* had a deletion instead of an insertion in subdomain 2 and *T. oceanica* Arp1 showed only three insertion, though this may be explained by the fact that this sequence was N-terminally truncated.

Arp1 is an essential component of the dynactin complex and the presence of a conserved Arp1 protein leads to the prediction of the presence of a dynactin complex in diatoms. Dynactin is a microtubule-associated protein complex involved in vesicular trafficking by modulating the motor protein dynein, that binds to cell organelles, which are to be transported along microtubules (Schroer, 2004). Screening the diatom genomes resulted in at least one homologue of a dynein heavy chain for all five diatom genomes. This confirmed and extended the results of Wickstead and Gull (2007) reporting dynein for *T. pseudonana* and *P. tricornutum*. The dynactin complex includes among other proteins a second member of the ARP family: Arp11 (Eckley and Schroer, 2003). Through BLAST search a sequence corresponding to Arp11 was identified in the *T. oceanica* proteome with a reasonable e-value of $1e^{-12}$. Another instance of Arp11 was discovered in the *T. pseudonana* genome which however, scored with only a moderate value of 0.001. Both proteins were not recognized as Arp11 by the ARPAnno web server, though an Arp11 homolog has previously been reported for *T. pseudonana* by Muller et al. (2005). Unfortunately no crystal structure of an Arp11 was available to conduct within this study. Though all members of the ARP family share the common clover leaf actin fold of the four subdomains and an overall sequence similarity with

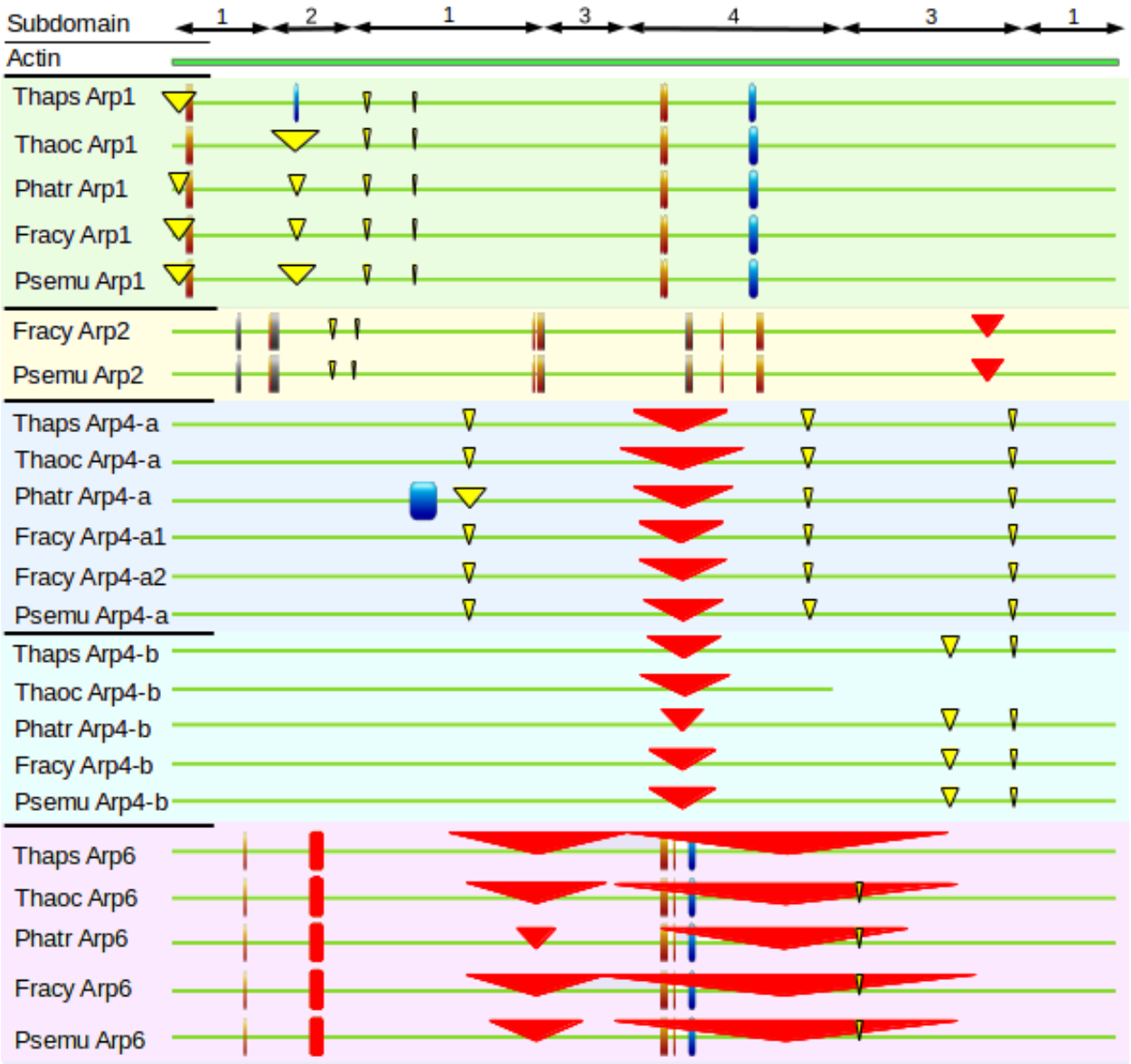


Figure 15.6: Graphically highlighted, simplified protein alignment of the diatom ARPs; sizes are drawn to scale. The green colored bar represents the 377 bp long actin sequence, black double arrows at the top, above the green line indicate the subdomains. ARP subfamily defining insertions, conserved throughout different phyla, are indicated in red triangles. ARP subfamily defining deletions are shown as red rectangles. Position of the insertions unique to diatoms, in comparison to the ARP subfamily classification are shown as yellow triangles. Deletions unique to diatoms are shown in blue rectangles. Discriminating residues for an ARP subfamily are colored in orange rectangles. Missing discriminating residues are shown in gray rectangles. Note that the Arp4 subfamily is divided into Group a and Group b.

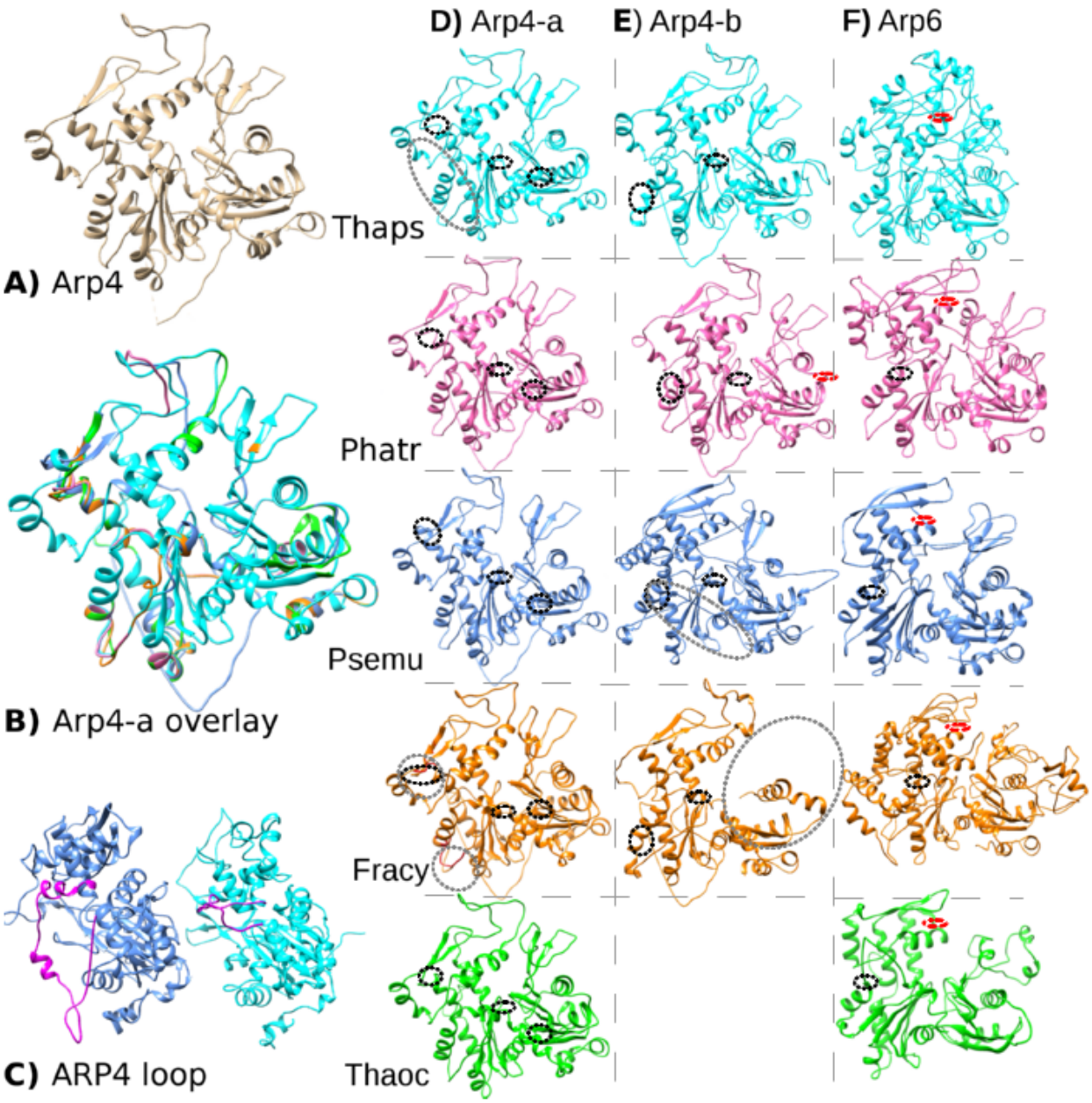


Figure 15.7: Results of the predicted tertiary structures of members in the diatom ARP subfamilies: Arp4-a, Arp4-b and Arp6. The ARP sequences are shown in the following color code: *T. pseudonana* — cyan, *P. tricornutum* — pink, *F. cylindrus* — orange, *P. multiseries* — blue, *T. oceanica* — green, and known crystal structures from other organisms are shown in gold. The diatom unique insertions (Fig. 15.6) are indicated in black dotted circles and the deletions in red dotted circles. Regions of special interest are shown in gray dotted circles. **A)** Crystal structure of *S. cerevisiae* Arp4 (3QB0). **B)** Superimposed structures of Thaps Arp4-a, Phatr Arp4-a, Psemu Arp4-a, Fracy Arp4-a1, Fracy Arp4-a2 and Thaoc 4-a. **C)** Characteristic loop of Arp 4 in pink.The rotated Psemu Arp4-b protein shows a turned loop and the rotated Thaps Arp4-a protein has a shortened loop in subdomain 3. **D)** Single tile representation of Arp4-a 3D-structures. A superimposition of Fracy Arp-4a1 (orange) and Fracy Arp-4a2 (red), showing high structural similarity with only two differences. For more detailed information about the shortened loop in subdomains 3 of Thaps Arp4-a (gray circles) see C). **E)** Tile display of 3D-structures of Thaps Arp4-b, Phatr Arp4-b, Psemu Arp4-b, Fracy Arp4-b. The Arp4-b fragment of Thaoc could not be calculated into a tertiary structure by Phyre[2]. Psemu Arp4-b has a turned loop (gray dotted circle) at subdomain 3, for more detailed information see C). The missing subdomain 2 and partly missing subdomain 1 of Fracy Arp4-b is due to a failure of the Phyre[2] software to predicted parts of the structure at the N-terminal portion of the protein. **F)**Tile display of 3D-structures of the superimposed Arp6. Subdomain 2 is only partially predicted by the Phyre[2] software.

actin, although they are generally found to be larger than actin (Poch and Winsor, 1997; Schroer and Fyrberg, 1994; Gordon and Sibley, 2005). Therefore the 3D protein structures of Thaps Arp11 and Thaoc Arp11 were superimposed with the crystal structure of an animal actin (*Oryctolagus cuniculus* - 3HBT), resulting in conserved areas especially in subdomains 1 and 3 (Fig. 15.2E). This lead to the conclusion that the two proteins are remote members of the Arp11 subfamily. The biochemical function of Arp11 in the dynactin complex is still unclear, but experiments in *Aspergillus nidulan* indicate, that Arp11 has a direct effect on the dynein–dynactin interaction by binding to the pointed end of the Arp1 filament of the dynactin complex (Zhang et al., 2008).

No homologues of Arp11 were detected in the genomes of the pennate diatoms. However since Arp1 and dynein were found in the present study in the pennate diatoms they should also possess the dynactin complex though without the Arp11 lock at the plus-end of the filament. It has been previously reported, that Arp11 is not an obligatory partner for Arp1 (Muller et al., 2005). Therefore the function of a plus-end-lock could potentially be provided by another plus-end binding protein such as formin, Aip1 or capping proteins. The presence of Arp1/Arp11 and consequently the putative existence of a dynactin complex is in marked contrast to the situation in red and green algae, as well as higher plants, where dynein dependent cytoplasmic organelle motility is not known (reviewed in Wickstead and Gull, 2007). Evolutionary diatoms are vastly distant from green algae and higher plants as they are chimeras originating from a secondary endosymbiosis between a unknown host and a unicellular red alga (Moustafa et al., 2009). It is very likely that the dynactin complex in diatoms originates from the ancestral non-photosynthetic host cell.

To all appearances, none of the diatom genomes contained a copy of Arp3, whereas a homologue of Arp2 was found and identified in two species, *F. cylindrus* and *P. multiseries*. Validated by the presence and absence of ARP-classification criteria using the ARPAnno web server the two sequences were predicted to be members of the Arp2 family with a score of 50.8 out of 100 (Fig. 15.5). This means that they were only slightly beneath the reliability threshold of 55. A detailed analysis showed that the two Arp2-like genes encode proteins which have 9 out of the 17 discriminating Arp2 motifs. Though this did include the highly conserved insertion in the C-terminal region of subdomain 3, other features again were not typical for Arp2 such as two short insertions in the N-terminal portions of subdomains 1 and 2 (Fig. 15.6). In addition other significant differences were detected in the ATP-binding site (Robinson et al., 2001). From the 10 highly conserved interacting sites of Arp2 with ATP (Asp^{158}, Asp^{161}, Gly^{162}, Val^{163}, Lys^{217}, Glu^{218}, Gly^{306}, Ser^{307}, Met^{309} and Tyr^{310}) only two residues Lys^{217} and Glu^{218} were conserved within the diatom Arp2s. These differences might be significant enough to compromise the ATP-binding and hydrolysis function of the diatom Arp2-like proteins. The predicted 3D structures of the diatom Arp2-like proteins were not as well conserved as the diatom Arp1 and Arp4 members in comparison to the experimentally determined crystal structures of these ARP subfamilies (Fig. 15.2 B, C, D). The absence of Arp2 from two diatom species and the unusual features of the Arp2-like proteins in the two remaining species as well as the absence of Arp3 altogether suggest, that diatoms do not contain the Arp2/3 complex. Further support for this suggestion came from the results of the BLAST search for other known protein components of the Arp2/3-complex such as C1 C2, C3, C4, C5. Non of these components were detected in the diatom genomes. Concluding from these results it can be assumed, that diatoms are incapable of inserting lateral branches at the actin filaments by means of the Arp2/3 complex. Diatoms therefore should feature a different mechanism of actin filament plus-end dynamics, that should have consequences for the structure and dynamics of the

actin networks in the diatom cytoplasm. The two tentative Arp2-like members identified in this study may rather serve a novel function, not related to actin filament branching.

Forming a branching network of actin filaments by means of the Arp2/3 complex enables animal cells to protrude their plasma membrane outwards such as in the case of lamellipodia (Borisy and Svitkina, 2000; Dominguez, 2010; Robinson et al., 2001; Pollard et al., 2000). Likewise, though more complicated through the presence of a cell wall, the Arp2/3 complex in plant cells is responsible for maintaining and shaping the cell boundaries (Mathur, 2005). Diatoms however, are encased by a stiff and inflexible silicified cell wall, the frustule. Considering the unique way how diatoms position the prefabricated silica frustule elements on the protoplasmic surface (Li and Volcani, 1985; Gordon et al., 2009) it appears conceivable that protrusion of cell boundaries is not a vital activity of cell morphogenesis in this group of organisms. Cellular morphogenesis of diatoms clearly does not involve gradual growth and reshaping of the cell surface as is the rule in most other organisms. However, complex nano-, meso- and micro-scale frustule patterning occurs within the SDV. It has been suggested, that the cytoskeleton plays a role in the shaping and remodeling of the SDV, thereby influencing the meso- and micro-scale patterning of the resulting silica wall structures (Pickett-Heaps and Kowalski, 1981; Schmid, 1987; Pickett-Heaps et al., 1990; Schmid, 1994; Van de Meene and Pickett-Heaps, 2002, 2004; Tesson and Hildebrand, 2010a,b). The mechanism responsible for this shaping process apparently does not involve the protrusion of cell boundaries by an Arp2/3 complex.

The lack of Arp2/3 is not unprecedented, as it has been reported from other unicellular eukaryotes such as Apicomplexa, Microsporidia and unicellular red algae (*C. merolae*) (Muller et al., 2005; Gordon and Sibley, 2005). The authors of these studies speculated that other related proteins such as Arp-like-protein-1 in Apicomplexa or formins take over the role of the ARP2/3-complex. Diatoms possess genes encoding formin proteins (see 15.3.3) that could be candidates for replacing the Arp2/3 complex. It is interesting to note, that the two pennate diatoms *F. cylindrus* and *P. multiseries* have a distantly related homologue of Arp2, whereas the centric diatom *T. pseudonana* and *T. oceanica* as well as the polymorphic pennate diatom *P. tricornutum* lack Arp2. The two diatom lineages (centric - pennate) diverged about 90 million years ago and a significant portion (40%) of the proteins encoded in the genomes are not shared between them (Bowler et al., 2008). As it is shown here, this even includes highly conserved protein families of fundamental function.

Another unique aspect of the diatom ARPs was the detected abundance and diversity of members of the Arp4 subfamily. BLAST searches within the five diatom genomes resulted a total of 11 different genes coding for Arp4 subfamily members. The ARPAnno web server clearly identified two Arp4 genes for each of the three species, *T. pseudonana*, *P. tricornutum* and *P. multiseries*, and three genes for *F. cylindrus* with a score of over 60 out of 100 (Fig. 15.5). *T. oceanica* also encodes an Arp4 protein with a score of over 60 and a second fragmented protein was detected which most likely is an Arp4 but could not be fully analyzed due to its C-terminal truncation (Fig. 15.6, Fig. 15.4). Typically, members of the Arp4 subfamily have about four insertions at conserved positions (Muller et al., 2005). On top of these subfamily defining insertions all diatom Arp4 members revealed a small unique C-Terminal insertion, which was conserved with respect to position and size. Sequence alignment (Fig. 15.6) showed additional insertions that caused the diatom Arp4 members to cluster into subgroup 4a and 4b. The two small insertions in subdomain 1 and 4 define subgroup 4a (Fig. 15.6). All diatom Arp4-a members showed a high 3D structural conservation (Fig. 15.7), even though their sequence identity was only 40% to 66%. The two

F. cylindrus Arp4-a members were very similar isoforms with an overall amino acid sequence identity of 92.5% (Fracy Arp4-a1 and Fracy Arp4-a2). Consequently, the 3D-structure was also very similar for these two isoforms (Fig. 15.7). Diatom Arp4-b proteins had one additional conserved insertion defining this group in subdomain 3 and were not as well structurally conserved within their group. The division into two subgroups is supported by the Maximum Likelihood tree (Fig. 15.4), with one representative of the subgroup for each diatom and in the case of Fracy-Arp4a even two. The apicomplexan comparative genome analysis of ARPs (Gordon and Sibley, 2005) revealed a similar split of the Arp4 subfamily into two groups, but unfortunately this feature has not been discussed by the author. Muller et al. (2005) remarked that within all observed phyla only two fungi *Y. lipolytica* and *S. pombe* encode a second Arp4 gene in their genomes, which he suggests is complementing the lack of Arp7.

Arp4 is present in the majority of all phyla and known to be a predominantly nuclear protein participating in chromatin-remodeling (Meagher et al., 2009; Dion et al., 2010; Fenn et al., 2011; Kast and Dominguez, 2011). Additional Arp4 is also involved in transcription regulation (Harata et al., 2002; Jiang and Stillman, 1996), DNA repair (van Attikum et al., 2004), histone acetylation (Galarneau et al., 2000) as well as kinetochore-spindle attachment, and gene silencing at centromeres (Minoda et al., 2005). Diatoms code for an above-average large amount of Arp4, that furthermore is clustering into two subgroups. This rises the question off the nuclear functions requesting Arp4, that are so important in diatoms compared to other organisms that only code for one isoform. It remains to be experimentally investigated, which roles the two Arp4 subgroups play in the diatoms. Whether they interact with different nuclear complexes, and what different functions they assume within the nucleus.

All screened diatom genomes encode one nuclear Arp6 protein recognizable by its complex insertion-deletion signature (Fig. 15.6). While this signature was highly conserved in the diatom Arp6s, except for one small insertion at the C-terminus which is missing in Thaps-Arp6 (Fig. 15.6), the overall structural conservation appeared to be weaker than in the other ARP-classes (Fig. 15.7 F). This might be due to the fact that although Arp6 is a ubiquitous member of the ARP family there is no crystal structure available yet to allow a fully reliable prediction of the three-dimensional structures. As to the function, Arp6 has been identified as a conserved subunit of the histone variant exchange complex (Meagher et al., 2009).

All diatom ARPs identified in this study have additional, diatom specific, small insertions (1-15 bp), which are not part of the classifying ARP-signature (Fig. 15.6, Tab. 15.3). Insertions are not unusual and define the ARP subfamilies. It has been shown across all eukaryotic phyla (algae, plants, protozoa, fungi, metazoans), that the insertion cluster at hotspots especially within subdomain 2, the smallest of the four subdomains, whereas subdomain 1 has the fewest insertions (Muller et al., 2005). In the diatoms, however, the hotspots are clustered in subdomain 1, whereas the smaller number of insertions occurs in subdomain 2 (Fig. 15.8). This feature may be regarded as the

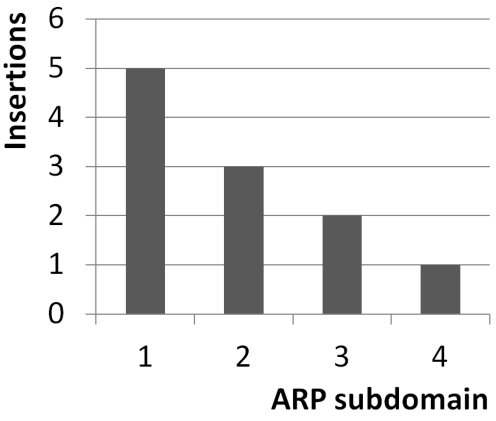


Figure 15.8: The sum of additional insertions occurring in the diatom ARPs missing from to the described and classified ARPs by Muller et al. (2005).

"diatom-signature" of the Arp-family genes.

Table 15.3: Discriminating motifs as well as position and size of the deletions and insertions of the diatom ARPs. DM —Discriminating motifs

		Insert								Deletion		DM
		1	2	3	4	5	6	7	8	1	2	
Arp1	Position Actin bp	5	50	79	94					50	231	
	Thaps Arp1	15	-1	2	1					1	2	5/5
	Thaoc Arp1	-	17	2	1					1	2	5/5
	Phatr Arp1	8	5	2	1						3	5/5
	Fracy Arp1	11	7	2	1						3	5/5
	Psemu Arp1	11	14	2	1						3	5/5
Arp2	Position Actin bp	64	73	323								
	Fracy Arp2	2	1	11								9/17
	Psemu Arp2	2	1	11								9/17
Arp4	Position Actin bp	42	96	119	203	253	264	311	334	96		
	Thaps Arp4-a	39	16	6	39	5	8		3			
	Thaoc Arp4-a	22	26	6	52	5	8		3			
	Phatr Arp4-a	18	-10	13	40	4	17		3	10		
	Fracy Arp4-a1	36		6	33	4	17		4			
	Fracy Arp4-a2	36	1	6	34	4	17		4			
	Psemu Arp4-a	18		6	31	5	21		4			
	Thaps Arp4-b	44	13		29		18	7	2			
	Thaoc Arp4-b	35			35		22	-	-			
	Phatr Arp4-b	34			17		81	7	2			
	Fracy Arp4-b	56	7		27		16	7	2			
	Psemu Arp4-b	23	9		27		129	7	2			
Arp6	Position Actin bp	147	246	275						56	207	
	Thaps Arp6	71	128	0						5	2	5/5
	Thaoc Arp6	15	134	2						5	2	5/5
	Phatr Arp6	16	98	2						5	2	5/5
	Fracy Arp6	55	153	2						5	2	5/5
	Psemu Arp6	36	137	2						5	2	5/5

Group	ARPs						Bundlers & Crosslinkers							Anchors to Membrans					Monomer Binders			Stabi-lisers	Depolymerizing & severing				Capping	
Proteins	Actin	Arp1	Arp2-like	Arp4	Arp6	Arp11	Alpha-actinin	Dematin-like	Fascin	Filamin	Fimbrin	Scruin	Villin	Annexin	Ankryn	Spectrin	Talin	Vinculin	Coronin	Profilin	Thymosin	Caldesmon	Tropomyosin	ADF/Cofilin	Aip	CAP	Severin	CapZ	Formin-like
Fracy	1	1	1	3	1	0	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	2	1	1	2
Psemu	1	1	1	2	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	2
Phatr	2	1	0	2	1	0	0	1	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	1	1	1	2	1	5
Thaps	1	1	0	2	1	1	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	1	1	1	1	0	6
Thaoc	1	1	0	2	1	1	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	1	1	1	1	1	7

Figure 15.9: Overview of the BLAST results for the screened ARPs and ABPs. Searches with no results are depicted as a black 0. Number of isoforms of the BLASTed proteins are shown in red. Proteins are clustered into groups, indicating their likely mechanism of action.

15.3 Actin-binding proteins in diatoms

15.3.1 Monomer binders & Depolymerizing proteins

BLAST searches revealed, that all five diatom genomes encode one copy of Aip1, ADF/cofilin, coronin and CAP protein each (Fig. 15.9), only *F. cylindrus* had an additional isoform of a CAP protein.

ADF/cofilin, Aip1, Coronin

The retrieved diatom ADF/cofilin proteins within this study were 140 amino acids long with an over all sequence identity of 62-89% between the five species. The *in silico* modeled tertiary protein structures of all diatom ADF/cofilin proteins showed the characteristic 3D structural folding into the conserved four anti-parallel central strands (β2-β5) and the two edge strands (β1 and β6) running parallel with the neighboring strands. The central mixed β-sheet was surrounded by two α-helices on each side (Fig. 15.10)(for a comparison see yeast cofilin Lappalainen et al., 1998). The conserved structural appearance is indicating, that the diatom ADF/cofilins are very presumably able to promote F-actin disassembly. ADF/cofilin have been proposed to be dominant proteins controlling F-actin dynamics by severing (Michelot et al., 2007).

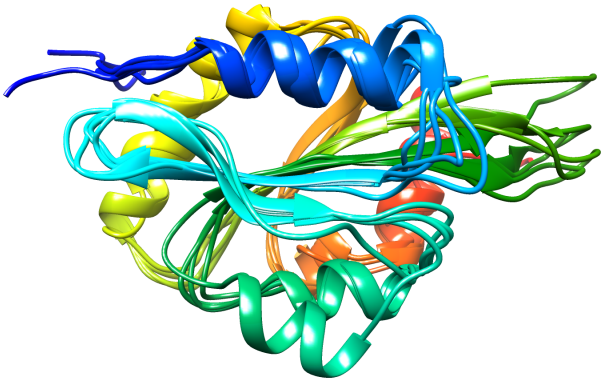


Figure 15.10: Resulting superimposition of predicted 3D structures of five diatom ADF/cofilin proteins. The tertiary structures were modeled and compared with the crystal structure of the yeast cofilin. All diatom ADF/cofilins into the central six-stranded mixed β-sheet, which is sandwiched between two pairs of α-helices. The protein backbone structure is depicted in rainbow colors from the N-terminus (blue) to the C-terminus (red).

The BLAST search for Aips resulted in genes coding for 660 to 720 amino acid long proteins, with predicted secondary structures of 9 to 11 WD40 repeats (Fig. 15.11 A). Although the di-

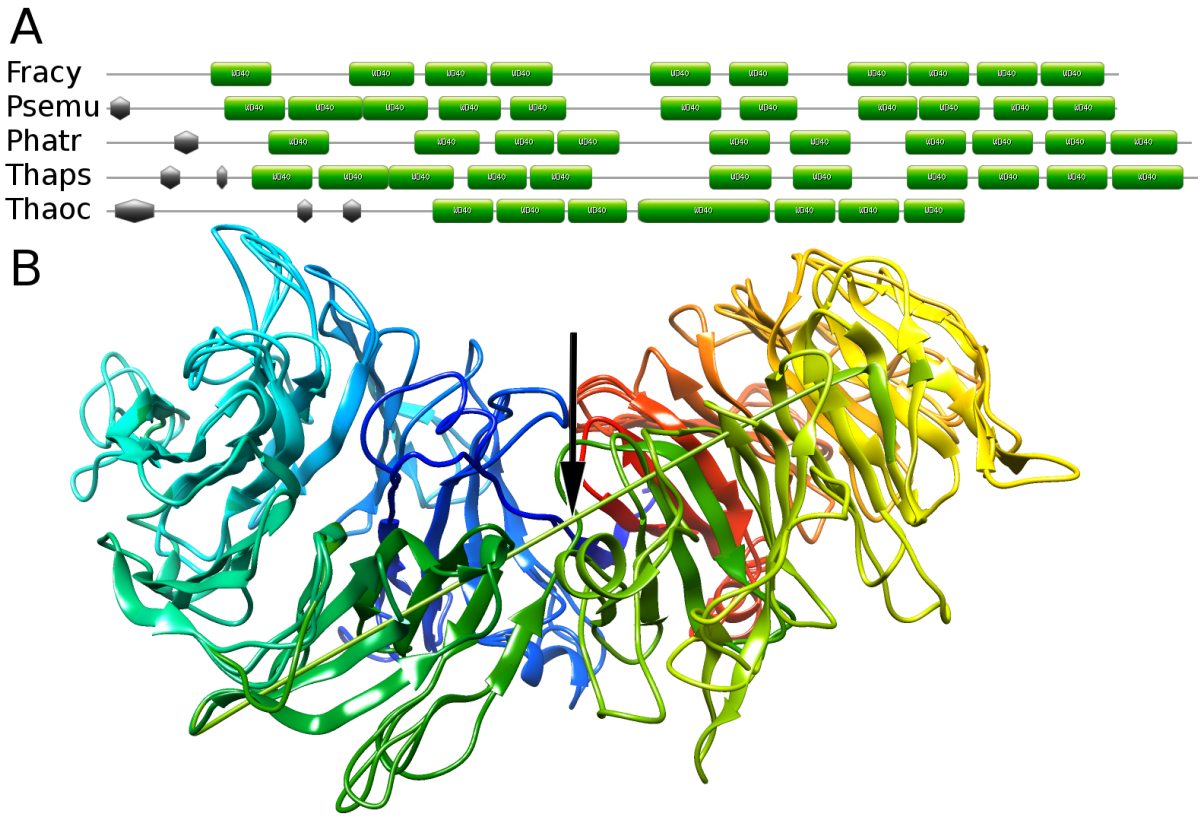


Figure 15.11: Results of the Aip1 sequence analyses. **A:** Aip1 secondary structure of the *F. cylindrus*, *P. multiseries*, *P. tricornutum*, *T. pseudonana* and *T. oceanica*. The *T. oceanica* Aip1 protein is truncated. Green rectangles — WD40 repeats, gray hexagon — region of low complexity. Secondary structure calculated by the SMART Software. **B:** Predicted 3D structure superimposition of the five diatom Aip1 proteins. The tertiary structure is quite conserved and the two β-propellers are separated by a bend when viewed from the side. The unpredicted region of *T. oceanica* is shown in a unfolded straight green line (black arrow). Proteins colored in rainbow colors from blue N-terminus to red C-terminus.

atom Aip1 amino acids identity varied between 23 - 76%, the computed 3D structures were quite conserved between the five proteins (Fig. 15.11 B), and folded into the typical two 7-blade WD40 β-propellers (Voegtli et al., 2003). This is in agreement with previous findings, that the secondary structure prediction of Aips generally result in a maximum of 10 WD repeats. Only after the crystal structure of Aip1 was resolved the native folding of Aip1 into a 7-blade WD40-double propeller was revealed (Voegtli et al., 2003). The Aip1 isoform of *T. oceanica* was an exception with only 7 WD40 repeats corresponding with an overall shorter sequence length, which is very probably due to incomplete sequence data. Nevertheless, the prediction of the 3D structure modeled the two β propellers, though there was an unpredicted region (Fig. 15.11 B black arrow).

The obtained diatom coronin sequences were predicted to fold into the typical coronin WD40-repeat containing β-propeller, including an additional diatoms unique 12 residue long insertion at amino acid position 113 (Fig. 15.12 A). Position, size, and residues of the insertion were conserved, indicating a functional importance of this motif (Fig. 15.12 B dotted circle). This insertion, not observed in other organisms, most likely has a diatom specific function. This function very probably is independent from the division into centrics and pennates 90 million years ago, because the sequence is even conserved between the two lineages. Bowler et al. (2008) observed previously, that considerable sequence diversity exists between centrics and the pennates and it would be therefore unlikely, that this highly conserved, structurally exposed motif has no function. The structures of the diatom coronins had the highest similarity to coronin 1a, that can bind to the plasma membranes (Gatfield et al., 2005). The N-terminal region of coronin 1a has a plasma membrane binding ability, localized in the same region as the diatom unique insertion. Therefore, diatom coronins could play a role as mediators between membranes and the actin cytoskeleton and it might be that the insertion influences the membrane binding activity.

The sequences, the secondary, and tertiary structures of the diatom ADF/cofilin, Aip1, and coronin proteins are all well conserved. Therefore, one can assume that disassembly through ADF/cofilin together with its enhancing and regulating partners Aip1 and coronin are fully functional within diatoms.

Profilin

Using different search strategies, i.e. BLAST, and profile hidden Markov models detected not a single gene coding for profilin in any of the diatom genomes. A possible candidate capable of replacing profilin, in its function as a G-actin binding protein, would be β-thymosin, a WH2 (WASP-homology 2) protein family currently known only in metazoans (Husson et al., 2010), but an extended search for this protein in the diatom genomes was also not successful (Fig.15.9). Though, several WH2 domain containing proteins were detected in diatoms, i.e. CAP described in chapter 15.3.1. In the present study a further organisms was found, that seemed not to contain a profilin gene: the enteric parasite *Blastocystis hominis* (Fig. 15.13). Due to its extremely reduced phenotype, its evolutionary origin has long been unknown. However, according to the rather recent molecular phylogenetic study of Yoshikawa et al. (2004), *B. hominis* is positioned at the base of the Stramenopiles clade together with the common marine zooflagellate *Cafeteria*, another colorless Stramenopila of which however, genomic informations are not yet available. In contrast within the genomes of other colorless Stramenopiles such as *Phytophthora sojae*, *Aurantiochytrium limacinum*, *Schizochytrium aggregatum*, and *Aplanochytrium kerguelense* profilin was detected, and the more

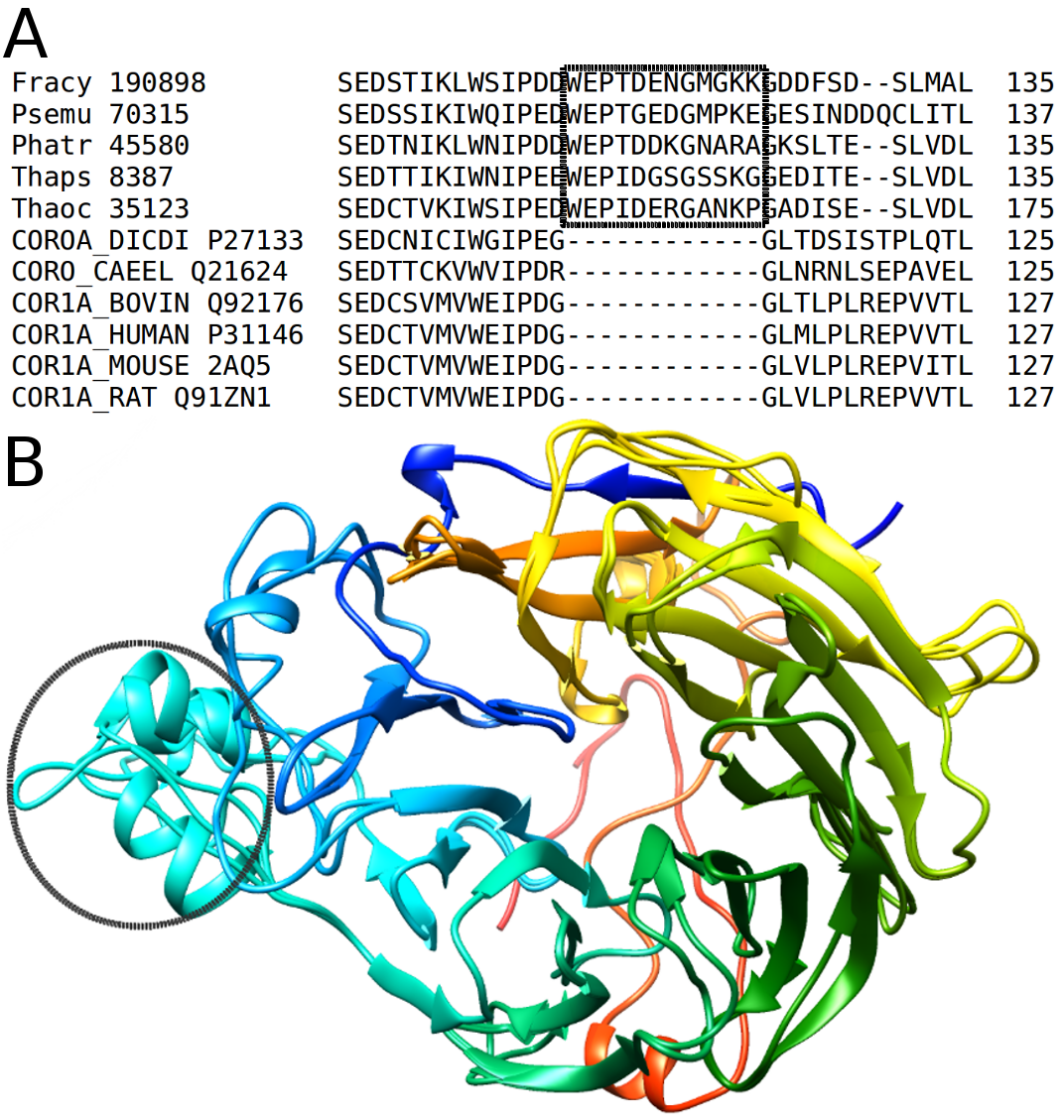


Figure 15.12: Results of the coronin sequence analyses. **A:** Partial sequence alignment of diatom coronins against reviewed UniProtKB coronin 1a protein sequences. All diatoms show a unique 12 aa long insertion (gray dotted rectangle) at amino acid position 113. **B:** Tertiary structure superimposition of the five highly conserved diatom coronins. The diatom unique, conserved insertion is a structurally exposed helix (dotted circle). DICDI — *Dictyostelium discoideum*; CAEEL — *Caenorhabditis elegans*; BOVIN — *Bos taurus*; HUMAN — *Homo sapiens*; MOUSE — *Mus musculus*; RAT — *Rattus norvegicus*. Protein backbone is depicted in rainbow colors.

closely related algal Stramenopila *Ectocarpus siliculosus* also encodes a profilin within its genome (Fig. 15.13). These profilins revealed classic features for this class of proteins, even though the overall homology with the two best matches identified by Phyre[2] (*Toxoplasma gondii* profilin, and *Schizosaccharomyces pombe* profilin) was only 34% .

So apparently, the ancestral Stramenopila host still had its copy of profilin, when it acquired a red algal symbiont to become the founder of the algal lineages of Stramenopiles, and the Phaeophyte clade maintained this copy, however, not so the diatoms. Considering, the red algal symbiont could have contributed another copy of profilin in its Stramenopila symbiosis the genome of the unicellular red alga *Cyanidioschyzon merolae* (Matsuzaki et al., 2004) was screened. A 197 amino acids long sequence (CMP220C), resulted to be a true version of profilin. Therefore, diatoms had potentially two separate origins for a profilin gene, yet apparently did not keep either of it. The finding that the Phaeophytes, which are a sister clade to the diatoms, exhibit a rather ordinary profilin, and

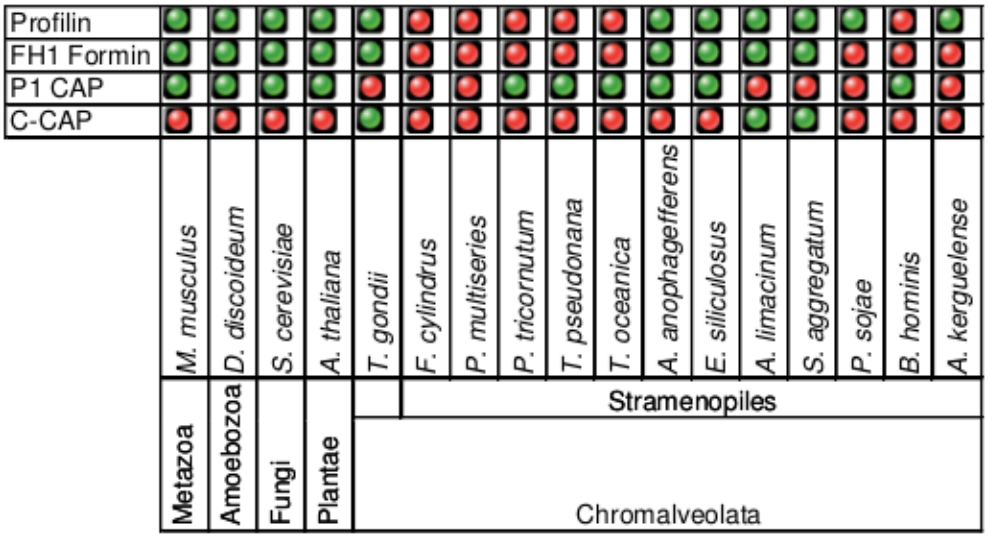


Figure 15.13: Presence of profilin, CAP containing the profilin binding site P1, and formins containing the profilin-binding site FH1 is listed for representatives of different eukaryotic phyla. Genomes encoding profilin, at least one CAP/formin gene including the P1/FH1 domains, are indicated in green circles. Genomes neither encoding profilin, nor CAP/formin containing the respective profilin-binding domains (P1/FH1) are indicated as red circles. *T. gondii*, *A. limacinum*, and *S. aggregatum* encode an N-terminally shortened C-CAP protein, and therefore do not have a P1 domain. Diatoms, as well as *B. hominis* do not encode profilin. The FH1 domain is missing from all diatom formins, and also from all *P. sojae*, *B. hominis*, and *A. kerguelense* formins. The P1 region of CAP is more widely distributed, only missing from two diatoms, and from *P. sojae*, and *A. kerguelense* .

given the fact that profilin serves important cellular functions (i.e. Cooley et al., 1992; Haarer and Lillie, 1990; Pollard and Quirk, 1994), it is even more intriguing, how diatoms can manage without it. Especially considering the extended networks of actin filaments that occur in both the pennales (chapter 7.1) as well as the centrales (data not shown). *B. hominis*, leading the life of an enteric parasite, apparently also can do without profilin, but clearly, both losses appear to have occurred independently of each other. Therefore probably another actin-binding protein functionally replaces profilin in diatoms.

Cyclase-associated proteins (CAPs)

CAPs consist of three functional defined regions, the N-terminal, the central, and the C-terminal domain (Hubberstey and Mottillo, 2002; Ono, 2013). Tertiary structure analyses of the N-terminal helically folded domain (HFD) of the diatom CAPs showed structurally highly conserved helices regions and less conserved coiled regions (Fig. 15.14 C). The C-terminal so called CARP domain of all diatom CAPs was also modeled into a very conserved 3D structure with regular β-sheet folding (Fig. 15.14 A,C). The conserved diatom HDF and CARP domain corresponded to the respectively domains on the CAP crystal structure (1k8F, Dodatko et al., 2004),indicating that both are fully functional in the screened diatoms. The HFD domain binds to ADF/cofilin-bound ADP-actin monomers, but functionally probability more important also to ADF/cofilin-bound actin filaments promoting severing (Chaudhry et al., 2013). Thereby the HFD domain resembles a function similar to Aip1. CARP, as well as WH2 in the central region of CAP are actin monomer binding sites promoting the nucleotide exchange of ADP–G-actin to ATP-G-actin (reviewed in Ono, 2013).

The center region of the CAPs is subdivided into two poly-proline stretches, P1 and P2, and a central WH2 domain. The P2 region was conserved in all six diatom CAPs as well as the WH2 domain (Fig. 15.14 A). The WH2 binds to G-actin at the same binding site as ADF/cofilin (Dominguez and Holmes, 2011), due to this feature WH2 is necessary for the nucleotide exchange

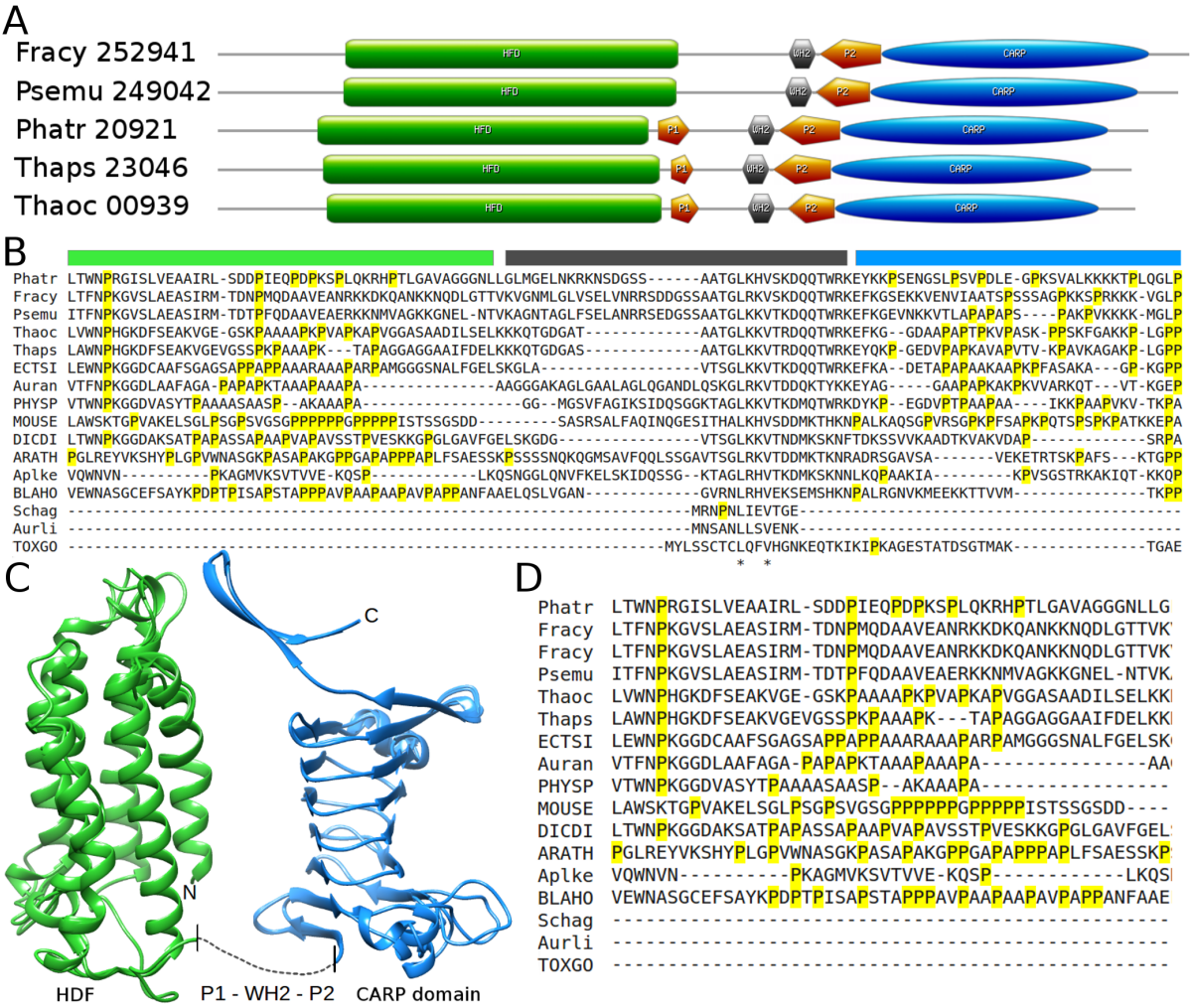


Figure 15.14: Graphical representation of the CAPs analyses. **A:** Schematic representation of the functional domains of the five diatom CAPs, drawn to scale. *F. cylindrus* is encoding an additional isoform of CAP which is very similar to Fracy 252941 and only slightly larger. *F. cylindrus* and *P. multiseries* CAPs contain no P1 region. HFD — green rectangles, P1 — right-sided orange pentagon; WH2 — gray hexagon; P2 — left-sided orange pentagon; CARP — blue oval. **B:** Sequence alignment of the five diatom CAPs in comparison with all screened Chromalveolata and one representative from plants, fungi, metazoa. Proline is highlighted in yellow; P1 —green beam; WH2 domain —gray beam; P2 blue beam. **C:** Superimposition of the tertiary structure of the HFD and CARP domains of the six diatom CAPs. There is no structure prediction for the central region (P1 - WH2 - P2), indicated as a gray dotted line. The orientation of the HFD and CARP domains to each other is unknown and was just chosen here for presentation. **D:** Enlarged sequence alignment of the P1 region. The polyproline region is defined by at least five prolines. No polyproline region could be detected for *F. cylindrus* , *P. multiseries* , *P. sojae* and *A. kerguelense*. Aligned sequence IDs: *P. tricornutum* Phatr 20921, *F. cylindrus* Fracy 252941, *F. cylindrus* Fracy 262057, *P. multiseries* Psemu 249042, *T. oceanica* Thaoc 00939, *T. pseudonana* Thaps 23046, *E. siliculosus* ECTSI 0124 0040, *Aureococcus anophagefferens* Auran 37435, *Phytophthora sojae* PHYSP 549399, *Mus musculus* MOUSE P40124, *Dictyostelium discoideum* DICDI P54654, *Arabidopsis thaliana* ARATH O65902, *A. kerguelense* Aplke 46244, *S. aggregatum* Schag 21210, *A. limacinum* Aurli 138808, *T. gondii* TOXGO 310030, *B. hominis* BLAHO D8M9H7.

when ADF/cofilin is bound to ADP–G-actin (Chaudhry et al., 2010; Nomura and Ono, 2013). The well conserved P1 region was detected in the CAPs of two centric diatoms and in *P. tricornutum* in a rather shortened version (5-6 residues), but absent altogether from the pennate diatoms *F. cylindrus* and *P. multiseries* (Fig. 15.14 A). In contrast, CAP alignments revealed that the P1 region consisting of at least up to 14 proline residues in other organisms, i.e. in *Mus musculus* (Fig. 15.14 D).

The P1 region is considered to be the profilin binding site (Bertling et al., 2007), and was either shortened or completely missing in the diatoms. Its function may be considered therefore as less stringent or even obsolete, which correlates well with the lack of profilin. The answer to the question why the two centrics and *P. tricornutum* still carry a residual polyproline stretch in the P1 domain, whereas the two other pennates do not, remains unknown. In contradiction to this result *B. hominis*, the only other known organism except for the diatoms not encoding a profilin gene, contained a CAP including a P1 region with 13 proline repeats (Fig. 15.13, Fig. 15.14 D). It can be therefore speculated that the loss of profilin occurred rather recently in *B. hominis* compared to the diatoms, as the P1 region within the *B. hominis* CAP was still strongly present. Further analyses of the Stramenopiles CAPs revealed, that *A. kerguelense* also lacked a P1 region (Fig. 15.13) and *P. sojae* CAP contained a truncated P1 region with only 4 proline residues (Fig. 15.14 D). The absence of the P1 region is an interesting co-evolution within the Stramenopiles and Apicomplexa CAPs, as it occurred independently in various species even though some species possessed profilin. The N-terminal truncated C-CAP proteins, described previously for Apicomplexa (Hliscs et al., 2010) were not only detected in *T. gondii* but also encoded in the genomes of *A. limacinum*, *S. aggregatum* (Fig. 15.14 B).

These results indicate a certain degree of variance in the mechanism of actin-subunit recycling in the Chromalveolata. Furthermore the Stramenopiles have a tendency to functionally replace profilin by CAPs. CAPs appear to be perfect candidates not only for ADF/cofilin recycling, but also for nucleotide exchange of monomeric actin subunits and G-actin sequestering. Diatoms very possibly make use of this mechanism, but likely also other Stramenopiles such as *P. sojae*, *A. kerguelense* and *B. hominis*.

15.3.2 Anchors to membranes

Screening the diatom genomes for genes coding for proteins (respective proteome for proteins) with the ability to anchor the actin cytoskeleton to membranes detected annexins and formins. Although some formins showed structural features with membrane binding abilities they are discussed in chapter 15.3.3.

Annexin

The genomes of the pennate diatom *P. tricornutum* and the centric diatom *T. pseudonana* both encode two annexin proteins including the four conserved annexin repeats. Analyzes of the sequences obtained from BLAST searches revealed that *T. oceanica* encodes one annexin with a truncated second annexin 2 repeat (annexin2) and an additional sequence fragment containing only one annexin repeat (Fig. 15.15). It remains to be seen, if this fragment belongs to a second complete annexin isoform. Homologous of annexin proteins were not found in the genomes of *F. cylindrus* and *P. multiseries*.

The annexins encoded in the diatom genomes showed the four characteristic annexin repeats

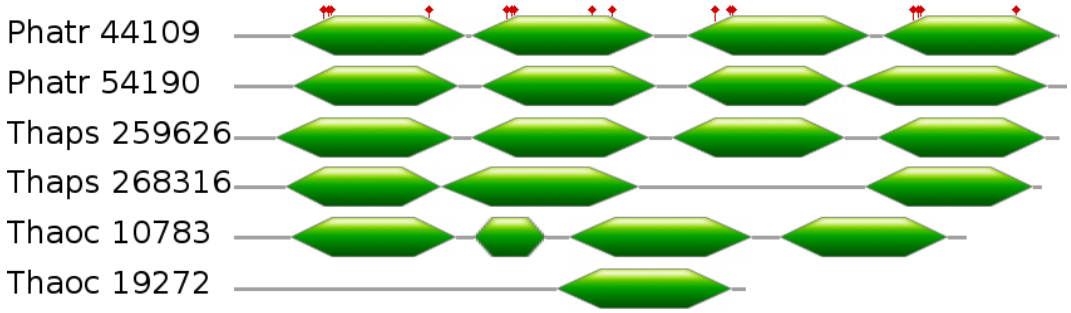


Figure 15.15: Calculated and graphical illustrated secondary structure of the five diatom annexins. Ca^{2+} binding sites (red flags) are shown exemplary for the *P. tricornutum* annexin. Position of the annexin repeats are indicated by the green double-pointed boxes. Note that the Annexin3 repeat of Thaps 268316 is not predicted by the InterPro or Smart software.

with a length between 61 aa and 80 aa (Fig. 15.15). Each of the repeats, expect the truncated *T. oceanica* annexin2 contained the characteristic 'type 2' Ca^{2+} binding motifs (GXGT-[38 residues]-D/E) (Barton et al., 1991). The secondary structures were quite conserved between the diatom annexins, although the sequence similarity among the members were only 42% - 63%.

The predicted tertiary structure was even more conserved between the diatom annexins (Fig. 15.16), leading to only a small conformational differences at the N-terminus as well as in proximity before and after the α-helix 11. The secondary structure of the missing annexin3 repeat and its 3D-structure of the *T. pseudonana* 268316 protein, were recalculated using the Phyre^2 software. This attempt resulted in a resolvable repeat with a truncated α-helix 12. Aligned diatom annexin sequences further supported the presence of the annexin3 repeat within Thaps 268316 and even revealed the residues for Ca^{2+} binding within this area. The annexin repeat2 in the Thaoc 10783 protein was truncated and α-helices 6-8 were missing.

The variability in the N-terminal domain of the annexin proteins could modulate given diatom annexins by their interaction and specificity to cellular target membranes. The N-terminal sequence of the diatom annexin proteins was rather short and varied between 17-25 amino acids. Short sequences in this range are known to be either unfold, or folded back onto the protein core (Liemann and Huber, 1997; Liemann and Lewit-Bentley, 1995). The actin binding motif IRI was only found in the Thaps 259626 protein. None of the diatom annexin proteins showed one of the other potential binding motifs for F-actin know from vertebrate and plant annexins LLYL-CGGDD and VLIRIMVSR (Hayes et al., 2004; Konopka-Postupolska, 2007). However, the protein domain responsible for annexin interaction with actin has still not been experimentally determined. Therefore it can not be finally concluded, if the diatom annexins apart form Thaps 259626 assume an actin-binding activity and anchor actin to membranes or not.

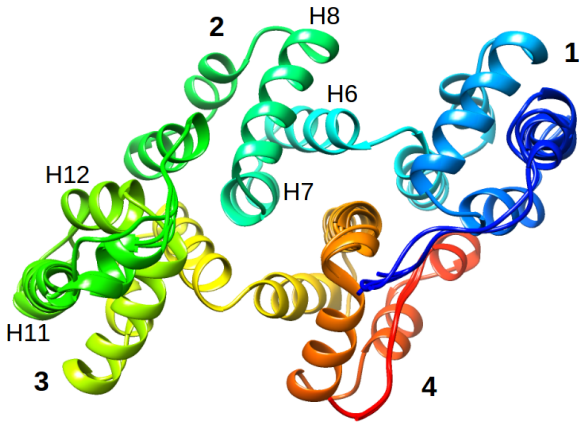


Figure 15.16: Superimposition of the predicted 3D structures of the five diatom annexins. Small conformational difference before and after the α-helix 11 can be observed as well as the short N-terminal domain. Numbers 1-4 indicate the four annexin repeats, α-helices of interest are indicated and numbered. Molecules are colored from blue — N terminus to red — C terminus.

15.3.3 Capping & Severing proteins

The representatives of capping proteins encoded in the five diatom genomes were CAPZ, severin and formin. No other capping proteins were detected in this study.

CAPZ

All screened diatom genomes, apart from *T. pseudonana*, encode an α- and β-subunit of the CAPZ heterodimer. Comparison of the diatom α-CAPZ sequences with α-CAPZ isoforms from various other species resulted in a high homology (Fig. 15.19 A). The core domains, consisting of five anti-parallel β-β sheets flanked left and right by two helices, were well conserved on the secondary and tertiary structural levels (Fig. 15.17 15.19 A). Whereas the N-terminal domain within the diatom α-CAPZs were found to be shortened and only a small set of 1-2 helices and 1-3 sheets were predicted within this region. The α-CAPZ sequence of *T. oceanica* was truncated presenting only the C-terminal fragment of 166 amino acids. Nevertheless it is most likely, that genome of *T. oceanica* contains a fully functional CAPZ heterodimer.

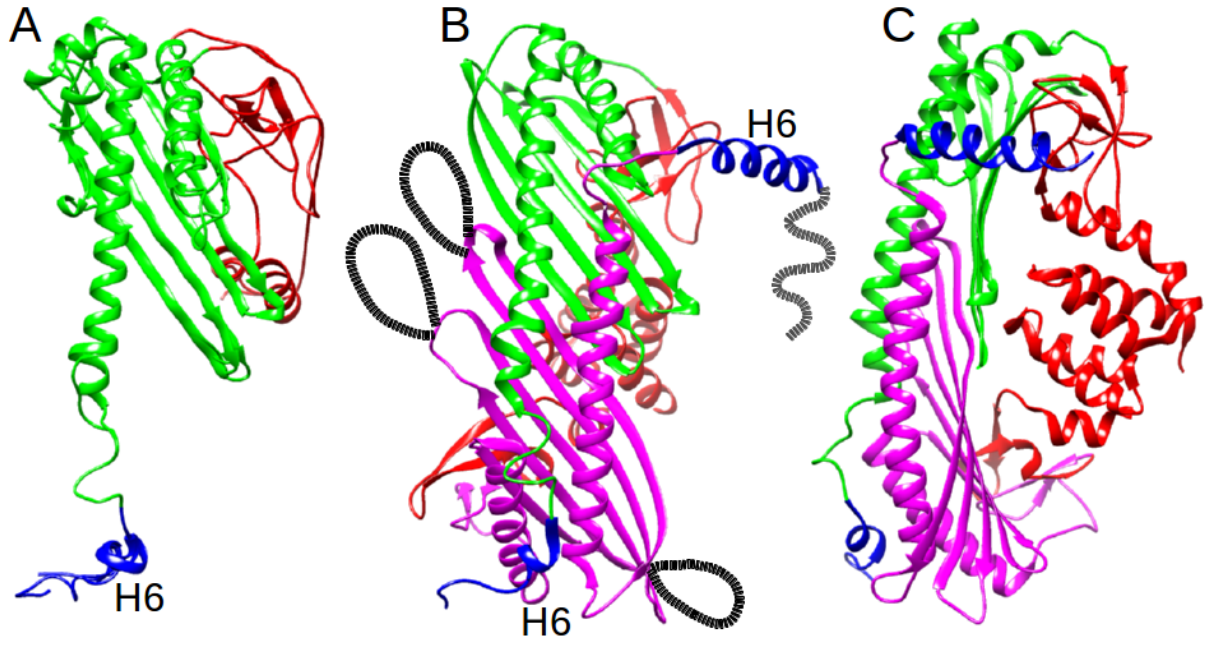


Figure 15.17: Results of the tertiary structural analyses of the diatom CAPZs. α-CAPZ — magenta; β-CAPZ — green; N-terminal — red; C-terminal — blue; Insertion — black dotted line; C-terminal extension — gray dotted line. **A:** α-CAPZ superimposition of the predicted 3D structure of *T. oceanica*, *P. tricornutum*, *P. multiseries* and *F. cylindrus*. Sowing the high conservation of the 3D structure within the diatom α-CAPZ. **B:** Crystal structure of *Gallus gallus* CAPZ heterodimer (1IZN:D). Unique insertion present within the core domain of *P. multiseries* and *F. cylindrus* are indicated in black dotted lines. The C-Terminal domain (H6 and flanking sequence) of the CAPZ protein signature crucial and exhibiting for actin binding activity are indicated in blue. The C-terminal extension of the *P. multiseries* and *F. cylindrus* β-CAPZ sequences is indicated in a gray dotted line. **C:** Side view of the crystal structure of the CAPZ. The N-terminal domain of the two subunits stabilizes the heterodimer formation.

The C-terminal domains including the helix 6 (H6) were conserved within the diatoms CAPZs with only one exception: *P. tricornutum* (Fig. 15.19 A). The H6 and especially the very C-terminal residues of the α- as well as β-CAPZ are known to be crucially required for actin binding (Hug et al., 1992; Casella and Torres, 1994; Sizonenko et al., 1996). It has been suggested, that within the CAPZ heterodimer each C-terminal end of the subunits binds to one of the two actin monomers

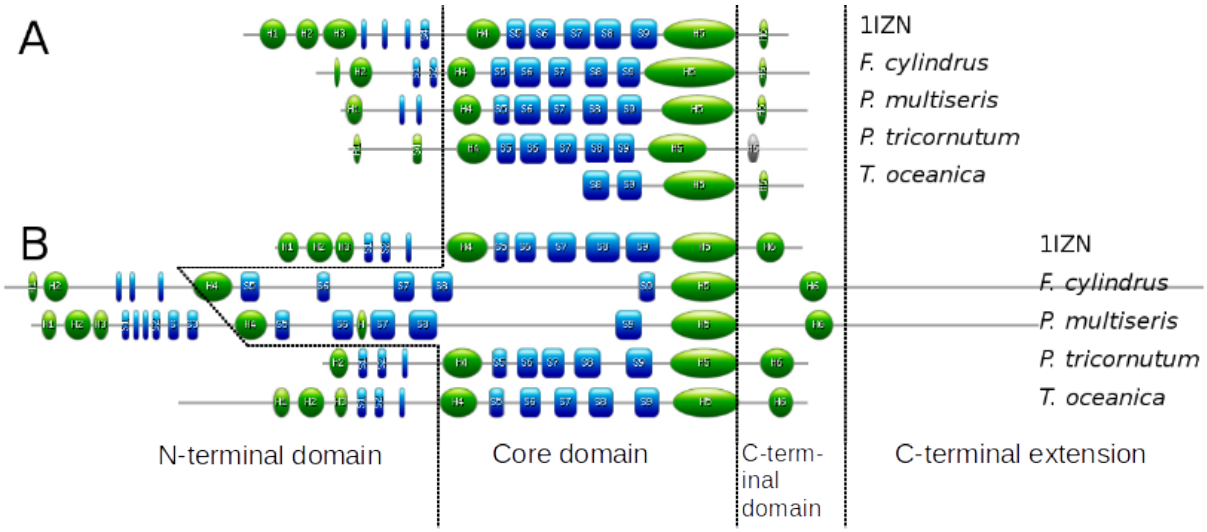


Figure 15.19: Graphical illustrated calculated and aligned secondary structure of the diatom CAPZ subunits in comparison with the *Gallus gallus* 1IZN structure: helices — green circles; sheets — blue rectangle. The three domains (N-terminal domain, core domain and C-terminal domain) are indicated by dashed lines. **A:** α-CAPZ. Only a fragment of the *T. oceanica* α-CAPZ protein was available. All diatoms show a reduced amount of sheets and helices within the N-terminal region in comparison to 1INZ. The Core domain is structurally very conserved as well as the C-terminal domain when compared to the crystal structure of 1IZN. *P. tricornutum* has a shortened C-terminal domain without H6. H6 and the very C-terminal domain of *P. tricornutum* as found in the EST database codon are indicated in light gray. **B:** β-CAPZ. *P. multiseries* and *F. cylindrus* have a long C-terminal extension (105 bp, 197 bp) and lager insertions in the Core domain in-between the anti-parallel sheets. Two of the insertions are conserved with respect to size and position. *P. multiseries* has two addition insertions within the N-terminal domain that are predicted to fold into sheets, whereas *F. cylindrus* has two additional N-terminal domain insertions with no structural prediction. Alignments were prepared with the Crustal Omega algorithm and drawn to scale in a simplified form.

at the plus-end of an actin filament (Yamashita et al., 2003) and thereby blocks the addition and dissociation of actin monomers.

From genomic sequence analyses it seemed, that *P. tricornutum* codes for a full C-terminal domain of α-CAPZ, which however is not transcribed due to the presence of a stop codon shortly after the first amino acid of H6. This would mean, that the CAPZ heterodimer of *P. tricornutum* with only one H6 (from β-CAPZ) probably is not able to bind to both, but only one actin monomer at the plus-end of F-actin. Therefore it might be possible, that the capping ability and strength of interaction with F-actin is reduced in the *P. tricornutum* CAPZ.

Screening the EST database of *P. tricornutum* with blastn and tbalstn for transcribed α-CAPZ sequences, revealed in the silica minus as well as in the silica plus database α-CAPZ sequences containing C-terminally elongated sequences that were highly similar to the H6 of *P. multiseries* and *F. cylindrus* (Fig. 15.18).

Calculating the secondary structure with PSIPRED, of the EST *P. tricornutum* α-CAPZ sequence including the C-terminal elongation, predicted an α-helix matching the size and position of H6 from α-CAPZs of other species (Fig. 15.19 A).

Therefore it seems convincing that *P. tricornutum* is encoding a fully functional α-CAPZ including the C-terminal H6. The stop codon within the genomic sequence of the α-CAPZ gene must therefore be consider as a sequencing artifact.

```
Fracy   W D A S A
Psemu   W D A N A
Phatr   W D V V A
        x x     x
```

Figure 15.18: H6 sequence obtained of the EST databases *P. tricornutum* silica plus-KS0ABA1YH13FM1.SCF:103-1069 compared with the sequence of H6 α-CAPZ sequence of *P. multiseries* and *F. cylindrus*.

T. oceanica and *P. tricornutum* β-subunit of the CAPZ (β-CAPZ) heterodimer were homologous to the β-CAPZs encoded in other organisms, although the N-terminal domain of the *P. tricornutum* β-CAPZ was shortened as well as reduced by two helices (Fig. 15.19B). An alignment of the *P. multiseries* and *F. cylindrus* β-CAPZ with β-CAPZs from the other species (*T. oceanica*, *P. tricornutum*, *Arabidopsis thaliana*, *Caenorhabditis elegans*, *Gallus gallus* and *Homo sapiens*) showed the clustering of the two proteins into the β-CAPZ signature and a long C-terminal extension of 105 respectively 197 amino acids (Fig. 15.19 B). This long C-terminal extension was unique to *P. multiseries* and *F. cylindrus* β-CAPZ. It is unclear what effect the C-terminal extensions may have, as screening for homologues of known structures and motifs did not lead to any results. For the function, there are two possibilities: either they might interfere with the actin binding ability of β-CAPZ by increasing/weakening the binding ability, or have some regulatory functions.

Furthermore, the gene analyses of *P. multiseries* and *F. cylindrus* β-CAPZs showed elongated coding sequences with unique, multiple insertions. Within the *F. cylindrus* β-CAPZ were five insertions: three of them in the core domain in-between the sheets and two in the N-terminal domain. *P. multiseries* had two larger insertions within the core domain and two insertions in the N-terminal domain. Two of these insertions within the core-domain of *P. multiseries* and *F. cylindrus* β-CAPZ were conserved with respect to position and size.

The insertions in the core domain lay in between the sheets (Fig. 15.17 B, C) and probably do not effect the core domain folding, therefore they are not likely to inhibit the heterodimer formation. The insertions in the N-terminal domain of the *F. cylindrus* and *P. multiseries* β-CAPZ proteins were predicted to fold into three additional sheets.

The N-terminal region of the *F. cylindrus* and *P. multiseries* α-CAPZ proteins are reduced in size and also in the number of helices and sheets, whereas the corresponding regions of the β-CAPZ proteins are increased. The N-terminal domain and the core domain are not directly interacting with the actin monomer and therefore do not change the F-actin binding ability. The N-terminal domain (Fig. 15.17 C) functions most likely as a stabilization of the heterodimer formation (Yamashita et al., 2003). Maybe the reduced α-CAPZ N-terminal domain is balanced by the increased N-terminal β-CAPZ domain of the *P. multiseries* and *F. cylindrus* CAPZ leading to a stable heterodimer formation. Whereas the shortened N-terminal domain of both subunits of *P. tricornutum* CAPZ (Fig. 15.19) are likely to influence the stability of the heterodimer.

F. cylindrus possessed a second copy of a β-CAPZ protein within its genome. This protein was not recognized by BLAST as a β-CAPZ protein, maybe because sequence informations were not fully available for the N-Terminal domain within the genome. Nevertheless the protein definitely showed the characteristic β-CAPZ signature confirmed by InterPro. The two β-CAPZ sequences encoded within the *F. cylindrus* genome had a nucleotide sequence identity of 77%.

No genes coding for a homologues of a CAPZ subunit were detected within the genome of *T. pseudonana*. Various search methods e.g. BLAST, PSI-BLAST, phmmer, and hmmsearch were conducted to strengthen this result. Even tBLASTn with an e-value of $1.0e^{-1}$ did not show any results for the mapped and unmapped sequences. Additionally no homologous to any CAPZ subunit sequence within the EST database of *T. pseudonana* could be detected. It was very unexpected, that all diatoms screened in this search code for one α-and β-CAPZ subunit each, except *T. pseudonana*. Especially since a species of the same genus encodes both subunits of CAPZ within its genome. It is very unlikely that both protein subunits of a heterodimer, that are

preserved in a wide variety of organisms and furthermore are also conserved in the Thalassiosira sister taxon, are missing in the genome of *T. pseudonana*. But it is also unlikely that the CAPZ sequences could not be detected by the search methods within the genome, because the sequences of CAPZ of various species are rather conserved. The assembled masked chromosomes still have multiple regions of up to 11.000 nucleotides with no sequence information within the completed genome of *T. pseudonana*. Maybe the homology sequences coding for CAPZ lay within this areas, although α- and β-CAPZ often do not even lay on the same chromosomes. It would be a very high coincidence if both subunits of CAPZ lay within areas of no sequence information. As CAPZ should be expressed throughout the cell cycle under various conditions but was absent from the EST databases it is after all possible that *T. pseudonana* does not encode a CAPZ heterodimer. CAPZ is the only capping proteins with no severing activities (Casella and Torres, 1994). Apart from the capping activity, CAPZ also stabilizes small actin oligomers leading the actin nucleation (Cooper and Pollard, 1985; Caldwell et al., 1989) and therefore assumes central functions within the actin regulation. Therefore it can be hypothesized that another actin-binding protein is functionally replacing the missing CAPZ in *T. pseudonana*, a potential candidate would be formin.

Formin

Most organisms encode at least two formins within their genome, some have even over a dozen formin-coding genes (Fig. 15.20). Within the genomes of the pennate diatoms *F. cylindrus* and *P. multiseries* a rather small number of only two formins were detected. *P. tricornutum* is an exception within the pennates and seemed to cluster more with the centrics, regarding the amount of formins encoded within their genomes. *P. tricornutum* encodes five, *T. pseudonana* six and *T. oceanica* seven homologous of formins. Two of these *P. tricornutum* formins were very similar isoforms, just slightly varying in size.

From the 22 obtained diatom formins ten were N- and/or C-terminally truncated (Tab. 15.5). The available databases were explored to achieve a complete inventory of the formin genes in their entire length. The C-terminally truncated sequences were elongated and translated until the next stop codon. For the N-terminal truncation the reason for truncation in all probability lay in the automatic gene model prediction from JGI and not in a true lack of this portion. Support from EST databases were not always available, as the proteins were rather large with up to 3700 amino acids and therefore were also often truncated within the EST databases although at different positions. When screening the nucleotide sequence of the N-terminally truncated proteins the next upstream stop codon followed by a start codon was for some gene models located even up to over 5000 bp upstream. The N-terminally truncated proteins were therefore elongated until the first occurrence of ATG after a stop codon, as suggested by Gruber and Kroth (2014). To map the diversity of formins within the Chromalveolata, Stramenopiles and Apicomplexa formin sequences were analyzed for known protein domains (Tab. 15.4).

Formins are described by the FH2 signature domain and most fomins also contain an FH1 domain almost invariably preceded N-terminal of the FH2 domain. Stramenopiles formin analyses clearly showed an evolutionary bias for the rare non-FH1 type formin genes as only 16 of the 70 Stramenopila formin genes contain an FH1 domain (Fig. 15.20). All diatoms as well as *B. hominis*, *A. kerguelense* and *P. sojae* exclusively encode non-FH1 formin genes (Fig. 15.20). The genomes of *A. limacinum* and *S. aggregatum* encode only one formin gene each with an FH1-FH2 architecture, but none the other encoded formin genes had a FH1 domain (Fig. 15.20). In contrast the two

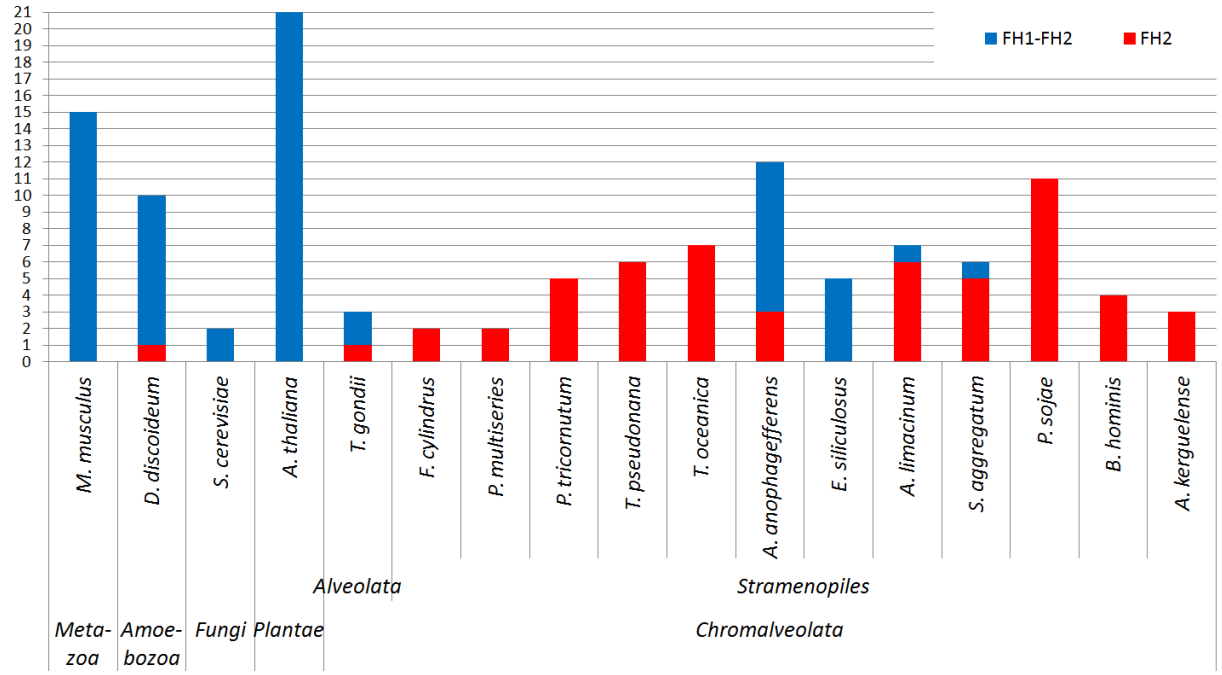


Figure 15.20: Formins encode in the genomes of representatives from different kingdoms. Formins containing the FH1-FH2 architecture and non-FH1 type formins are listed. In total 70 Stramenopila formins have been analyzed. The Chromalveolata especially the Stramenopiles are observed in more detail. Only within the Stramenopiles are species that encode exclusively formins with no FH1 domain. In general Chromalveolata seem to have more genes encoding formin genes with a FH2 architecture, than formins with a FH1-FH2 architecture.

screened non-diatom phototrophs *A. anophagefferens* and *E. siliculosus* had their majority of genes including an FH1 domain. Most notable was *E. siliculosus*, where all the formins had an FH1-FH2 architecture. Outside the group of Stramenopiles a missing FH1 domain is a rare event. The *D. discoideum* formin ForC and the *Plasmodium* formin (MISFIT) belonging to the Apicomplexa are the only known none Stramenopiles formins that also lack the FH1 domain (Kitayama, 2003; Higgs and Peterson, 2005; Rivero et al., 2005; Bushell et al., 2009). Unfortunately the functions of the two non-FH1 type formins could not be fully clarified by the authors. In contrast to the diatoms, however, *D. discoideum* has several more formins including the FH1-FH2 signature. Grunt et al. (2008) previously observed, that *T. pseudonana*, *P. tricornutum*, *Phytophthora ramorum*, and *Plasmodium falciparum* possess exclusively formins without an FH1 domain. Extending the screening in the present study for further Apicomplexa genomes resulted that *T. gondii* and *Cryptosporidium parvum* are encoding at least one formin containing an FH1 domain and that *Plasmodium falciparum* is an exception within the Apicomplexa exclusively possessing non-FH1 formin genes. As far as diatoms are concerned, the results of Grunt et al. (2008) could be confirmed and extended in the present approach by three additional diatom species, suggesting that the entire phylum of Bacillariophyta is characterized by non-FH1-type of formins.

The function of the FH2 and FH1 domain of formins is well studied. The FH2 domain is known to mediate homodimer formation and as a homodimer formin is capable to bind to the plus-end of the actin filament (reviewed in Paul and Pollard, 2009). On the other hand, the FH1 domain is known to interact with profilin-G-actin (Courtemanche and Pollard, 2012). The interaction of the FH1 domain with profilin-G-actin is leading to an acceleration of the plus-end polymerization. Diatoms have no profilin (see 15.3.1), therefore it is most likely, that in the course of evolution the FH1 domain has been lost from the formin genes. As previously discussed for the CAP P1

domain (see 15.3.1), the loss of the FH1 domain within the Stramenopiles in all likelihood was an independent event. Profilin is not encoded in both the diatoms and Blastocystis (Fig. 15.13), therefore the independent FH1 lost might possibly result from the lack of functional selective pressure on that domain.

The lack of the FH1 domain in the diatom formins very probably lowers the kinetics of actin filament plus-end polymerization, leading to the consideration that these formins may have a capping function. The formin homodimer at the plus-end of a filament has been previously described to act as a *leaky cap* that allows elongation of the filament, while capping it at the same time (Zigmond et al., 2003). Previous experimental data showed, that the presence of the FH1 domain switches the FH2 domain in its dimer configuration from a tight capper to a *leaky* capper (Michelot et al., 2005), leading to the conclusion that the diatom formin proteins might indeed be capable of efficiently capping F-actin. In that context it is noteworthy that *T. oceanica* does not encoding any capping proteins and all the other screened diatoms encode only one CAPZ each (see 15.3.1).

Furthermore, formins are described to have the ability to nucleate actin filament in a Arp2/3-independent pathway. The isolated FH2 domain of the wild-type formin (Bni1p) showed a nucleation activity of 50%, when compared to the reactivity of Bni1p (Pruyne et al., 2002). Apparently, the regions N-terminal of the FH1 domain as well as C-terminal of the FH2 domain do not seem to have any functional significance in actin nucleation (Pruyne et al., 2002). Diatoms do not possess an Arp2/3 complex (see 15.2), it is therefore a very likely possibility that formins assume this function, although formins nucleate straight filaments and the nucleation of the Arp2/3 complex leads to a branched network.

Formins are essential for the assembly for polarized arrays of actin cables (Sagot et al., 2002a), but the mechanism behind it is not fully understood (reviewed in Blanchoin et al., 2014). After nucleating actin filaments they can move to the lateral side of an actin filament, where they nucleate another actin filament extending parallel to the existing one and therefore effectively build F-actin bundles (Michelot et al., 2006; Esue et al., 2008). The formins with a deleted FH1 domain however, have very weak bundling activity, leading Michelot et al. (2005) to the conclusion that FH1 is essential for the bundling process. Taking this into account diatom fomins are probably capable of bundling actin cables, although to a much lesser extend.

Formins are known to be able to anchor actin filaments to membranes (reviewed in Cvrcková et al., 2004) and analyzing the domain architecture of Stramenopila formins revealed frequently membrane binding motifs (Tab. 15.4). Within the genus Thalassiosira, *T. pseudonana* possessed three formin genes with the ability to bind to membranes, whereas *T. oceanica* formins had no obvious membrane interacting motifs (Fig. 15.21). However, six of the seven *T. oceanica* formins were N-terminal truncated, it therefore seems to be quite likely, that the genome of *T. oceanica* codes for formins containing membrane binding domains N-terminal of the truncation.

Within the genome of the centric diatom *T. pseudonana* two formin genes (Thaps 9248, 24787) contained a Pleckstrin homology domain (PH domain) that can bind Phosphatidylinositol lipids within biological membranes (Fig. 15.21). The predicted PH-domain, known to be involved in recruiting proteins to different membranes (Wang and Shaw, 1995), could very probably target diatom formins to appropriate membranes of cellular compartments.

The third *T. pseudonana* formin homolog (Thaps 8340) contained a Phosphatase tensin-type domain (PTEN) as well as a tensin-type C2 domain (C2-TENSIN), that has the ability to bind

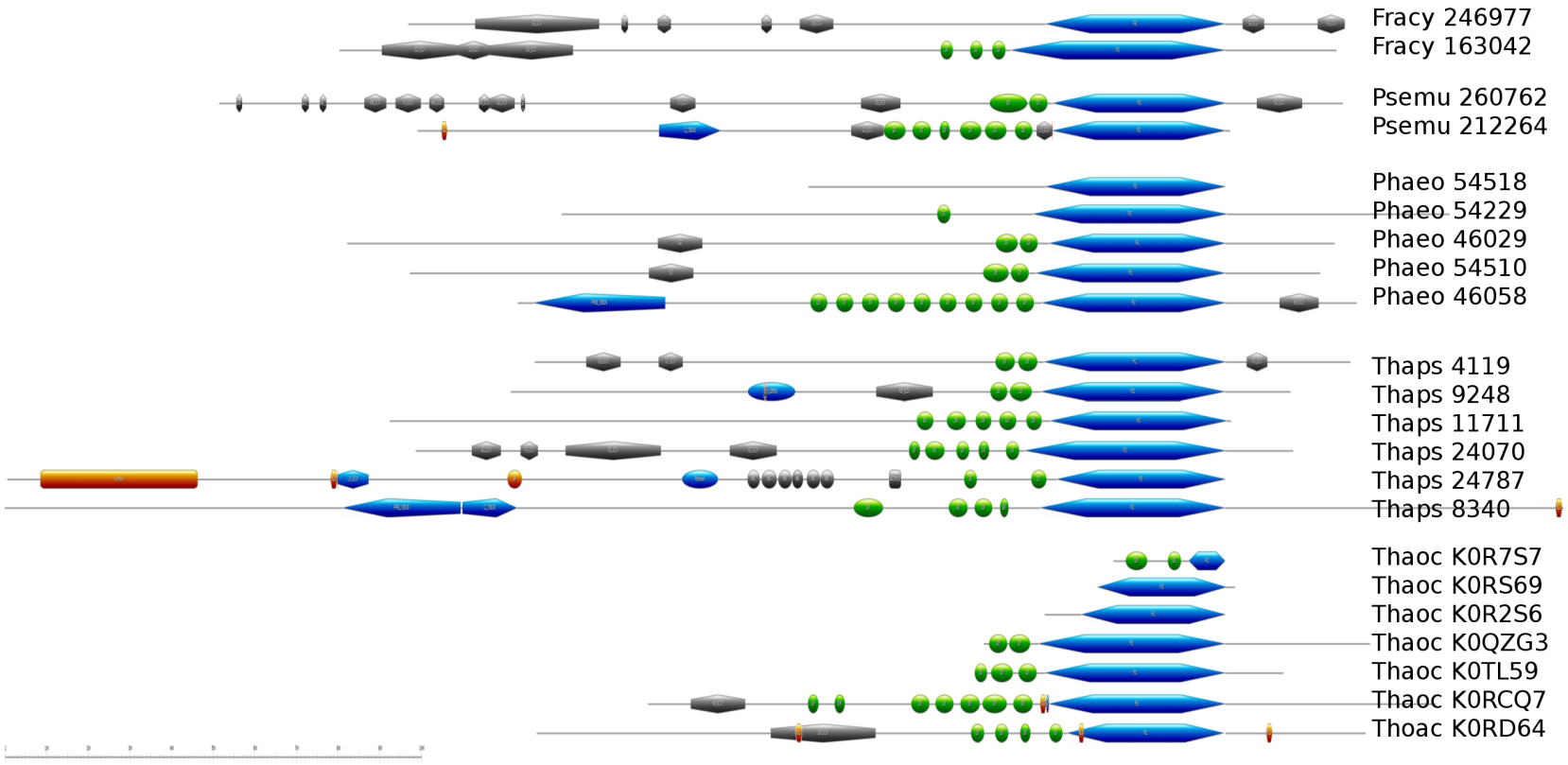


Figure 15.21: Graphically illustrated predicted secondary structure of the diatom formins clustered by species. All diatom formins have the conserved FH2 domain (blue hexagon), the FH1 domain is missing from all diatom formins. DUF2360 (green circle) are regularly encoded within the formins. Psemu 212264, Thaps 24787/8304 and Thaoc K0RCQ7/K0RD64 have NLSs (orange hexagon). Membrane binding domains (blue circle — PH domain; blue right pointed heptagon — CTENSIN; blue left pointed heptagon — PTEN) are within the Psemu 212264, Thaps 9248/24787/8340. Two of the *P. tricornutum* formins (Phaeo 46029, 54510) share a 98.58% identity on the amino acid level. Orange circle — SAP, gray circle — ankyrin-repeats, gray hexagon — regions of repeated amino acids. Drawn to scale, scale bar 1000 aa.

Table 15.4: Protein domains predicted of Stramenopila formin gene sequences using SIB Motif Scan. Additional not listed domains are: Thaps: SAP, KIF4; *A. anophagefferens*: UVRD-HELICASE-ATP-BIND, DNA2NAM7, SRP72, PHOSPHOPANTETHEINE, EGF, PT, REJ, ARM-Repeat, EF-HAND, CNMP-BINDING-3, HEMOPEXIN, TTL; *A. limacinum*: Allatostatin; *S. aggregatum*: CARBAMOYLTRANSFERASE; *P. sojae*: RUN.

	FH1	GBD FH3	DUF	PH	C2	ANK	ZF-Fyve	P-TEN	NLS	TPR	WW	ZF-PHD
T. gondii	+											
F. cylindrus			+									
P. multiseries			+		+							
P. tricornutum			+					+				
T. pseudonana			+	+	+	+		+	+		+	
T. oceanica			+						+			
A. anophagefferens	+	+	+	+					+			
E. siliculosus	+	+	+						+			+
A. limacinum	+	+	+		+	+	+	+	+	+	+	
S. aggregatum	+	+	+	+	+	+	+		+	+	+	
P. sojae		+	+				+		+			
B. hominis		+										+
A. kerguelense		+	+							+		

phospholipid membranes (Lee et al., 1999). *P. multiseries* and *P. tricornutum* both encode one formin gene with clear membrane interacting domains: Psemu 212264 with an C2-TENSIN domain and Phaeo 46058 with an PTEN and by the analyzing software weakly supported an additionally C2-TENSIN domain (Fig. 15.21). Within the Stramenopiles *A. limacinum* also encodes a formin gene containing a C2-TENSIN domain (Tab. 15.4). The diatom C2-TENSIN formins potentially anchor actin filaments to membranes and as formins are also actin nucleators they might therefore enable actin nucleation at membranes in diatoms. The C2-TENSIN domain has not been detected in plants so far (Cvrcková et al., 2004) and seems to be rather restricted to the Stramenopiles. The PTEN domain can lead to peripheral association with membranes (Cvrcková et al., 2004) as shown for two plant class-II formins (Zhang et al., 2011; van Gisbergen et al., 2012). Grunt et al. (2008) claims that the PTEN-FH2 architecture of formins in plants and Stramenopiles are of independent origin.

Notably, the genome of *F. cylindrus* did not encode any formins with domains known to interact with membranes, whereas probably all other diatoms do. Since this is very unlikely, a more extensive sequence analysis should be performed to find hidden motifs and it should also be considered that membrane binding could be mediated via protein-protein interaction domains between forming and an integral or peripheral membrane proteins.

The two centric diatoms encode two formins each, containing a nucleus localization signal (NLS) (Fig. 15.21, 15.5). One *T. oceanica* formin (Thaoc K0RD64) even showed up to 3 NLSs. Within all pennate diatoms only one forming was found in *P. multiseries* that contains an NLS supported by a low e-value. Other NLS motifs were detected using Motif Scan of ExPASy in two

P. tricornutum formins and two *F. cylindrus* fomins and one *P. multiseries* formin, although all with a not reliable high e value of $2.1e^{+04}$. Judging from the NLSs within the diatom formins it seems convincing that diatoms encode formins, that are involved in the regulation of the nuclear actin. Unexpected was the differentiation between the pennate and centric diatoms, where within the centrics two formins contained NLSs and within the pennates only one *P. multiseries* formin contained a NLS. One can speculate that within the centrics the regulation of the nuclear actin is more depending on formins, or that the recruiting of fomins into the nucleus is more depending on the NLS in the centrics than within the pennates.

Although *F. cylindrus* and *P. tricornutum* do not encode formins with a defined NLS, they still had at least one formin each with a potential NLS even though only supported by a high e value. It is still not fully understood how the highly dynamic shuttling of formins between the cytoplasm and the nucleus is regulated (reviewed in Baarlink and Grosse, 2014). Therefore it is also still possible, that these formins are also involved in the nuclear actin regulation, containing up to now unknown signals to get recruited to the nucleus.

Formins within the nucleus are considered to be associated with the inner nuclear envelop (Baarlink and Grosse, 2014). This assumption would coincide well with the gene models Thaps 24787/8304, Psemu 212264 and Thaoc K0RCQ7. Thaps 24787/8304 and Psemu 212264 which are capable to interact with membranes due to the C2-TENSIN, PTEN and PH domains and additionally contain at least one NLS. Although *T. oceanica* did not show any obvious membrane interacting domains, Thaoc KORCQ7 as well as Thaps 24787 had a Lysin-Rich region right next to the NLS. It has been experimentally proven for the formin mDia2 that a basic stretch of amino acids just next to the NLS mediates attachment of mDia2 to lipid membranes (Gorelik et al., 2011; Ramalingam et al., 2010). Applying this information to the Thaoc KORCQ7 sequence, it is likely that this formin is also able to attach to the inner membrane of the nuclear envelop, although it has no membrane interacting domain. In conclusion, it is reasonable to propose that centric diatoms and *P. multiseries* have formins that shuttle between the cytoplasm and the nucleus and are able to associate with the inner nuclear envelop.

Table 15.5: Domains and regions of the diatom formin proteins. The domains and regions are listed in the same order as occurring in the proteins. Truncated N-terminal (N) or C-terminal (C) sequences were manual elongated (ME), in the case were sequence information was available. Abbreviations: PTEN — Phosphatase tensin-type domain profile; C2-TENSIN — C2 tensin-type domain profile; CHITINASE-18 — Chitinases family 18 active site; Kinesin — Kinesin motor domain; SAP — SAP motif profile; PTB — Phosphotyrosine-binding domain; ANK-REPEAT — Ankyrin repeat profile; WW-DOMAIN-1 — WW/rsp5/WWP domain signature; PH-DOMAIN — Pleckstrin-homology domain; ALD-DEHYDR-GLU — Aldehyde dehydrogenases glutamic acid active site; NLS — Bipartite nuclear localization signal profile; ALA-RICH — Alanine-rich region profile; ARG-RICH — Arginine-rich region profile; ASP-RICH — Aspartic acid-rich region profile; GLN-RICH — Glutamine-rich region profile; GLU-RICH — Glutamic acid-rich region profile; GLY-RICH — Glycine-rich region profile; LYS-RICH — Lysine-rich region profile; PRO-RICH —Proline-rich region profile; SER-RICH — Serine-rich region profile; THR-RICH — Threonine-rich region profile.

Sequence ID	Length	Domain, Regions	Truncated	ME

Thaps 24070	2,107	SER-RICH, THR-RICH, 2 x SER-RICH, FH2, 5 x DUF		
Thaps 8340	3,742	PTEN,C2-TENSIN, 4 x DUF, FH2, NLS		
Thaps 24787	2,910	Kinesin, ,LYS-RICH, NLS, SAP, PH domain, 6 x ANK, WW DOMAIN, 2 x DUF, FH2	N	Yes
Thaps 11711	2,020	5 x DUF, FH2		
Thaps 9248	1,873	PH-DOMAIN, ALD-DEHYDR-GLU, ASP-RICH, 2 x DUF, FH2		
Thaps 4119	1,960	GLN-RICH, GLY-RICH, 2 x DUF, FH2, PRO-RICH		
Thaoc K0R7S7	269	2 x DUF, FH2	C/N	
Thaoc K0RS69	330	FH2	N	
Thaoc K0R2S6	434	FH2	C/N	
Thaoc K0QZG3	928	2 x DUF, ARG-RICH, FH2	N	
Thaoc K0RD64	1,990	SER-RICH, NLS, 4 x DUF, FH2, 2 x NLS	N	
Thaoc K0TL59	742	3 x DUF, FH2	N	
Thaoc K0RCQ7	1,892	ALA-RICH, 7 x DUF, LYS-RICH, NLS, FH2		
Phatr 46058	2,015	9 x DUF, FH2		
Phatr 54229	2,132	DUF, FH2		
Phatr 54518	1,001	FH2		
Phatr 46029	2,371	SER-RICH, DUF, FH2		
Phatr 54510	2,186	SER-RICH, DUF, FH2		
Psemu 260762	1,169	SER-RICH, 2X DUF,Chitinases, FH2, SER-RICH	C/N	Yes
Psemu 212264	1,591	NLS, C2-TENSIN,GLU-RICH, 6x DUF,LYS-RICH, FH2	N	Yes
Fracy 246977	2250	SER-RICH, ASN-RICH,FH2, GLN-RICH, 2xGLN-RICH, SER-RICH, THR-RICH		
Fracy 163042	2394	SER-RICH, GLN-RICH, 3 x DUF, FH2	C-/N	Yes

A very specialized formin with an exceptional domain pattern was Thaps 24787. Thaps 24787 contained six ankyrin repeat domains, one WW-domain, one PH-domain, one NLS and one SAP motif, and on top of it one N-terminal kinesin motor domain (Fig. 15.21). The WW-domain as well as the ankyrin repeat domains are interaction sites for proteins. Whereas the SAP motif is a putative DNA-binding motif named after individual nuclear proteins: SAF-A/B, Acinus and PIAS

(Aravind and Koonin, 2000). SAP is present in a variety of nuclear proteins involved in transcription, DNA repair, RNA processing, or apoptotic chromatin degradation. BLAST results verified the kinesin motor domain with an e-value of $4e^{-81}$ to be a chromosome-associated kinesin KIF4. The combination of the above mentioned domains including the KIF4-motor domain makes this formin a likely candidate to assume functional roles during mitosis.

Up to now actin has been discussed to play a role during the spindle positioning and formation, furthermore there is a strong crosstalk between cytoskeletal actin and mitosis, but the role of actin within the spindle still lacks strong evidence (reviewed in Lancaster and Baum, 2014). The motif architecture of Thaps 24787 formin gives good evidence of a connection between the nuclear actin and the microtubules and therefore would be a very interesting protein to study *in vivo* to observe its function and role through mitoses.

Nearly all of the analyzed diatom formins had multiple copies of domains of unknown function (DUF2360). The DUF2360 repeats were all N-terminal of the FH2 domain and varied from one up to nine repetitions (Fig. 15.21). Phaeo 54518 was the only formin lacking DUFs, with fully available sequence information. The gene model of Fracy 246977 also showed no clear DUF domain N-terminal of FH2, but within this area lay an unsequenced region of unknown length. Same applied for the two *T. oceanica* formins (Thaoc K0RS69, K0R2S6) which were only C-terminal fragments breaking off right in front of the FH2 domain and were therefore not taken into account. The multiple DUF2360 motif was widespread within the Stramenopiles, only *B. hominis* formins lacked any DUF domain (Tab. 15.4). The regular expression of the multiple DUF region within all screened Stramenopila formins is indicating a region of importance even though the function of DUF is unknown.

A well studied formin regulatory domain is the GBD-FH3 motif: a widespread though not ubiquitous motif, regulated by Rho-family GTPases (Paul and Pollard, 2009). Grunt et al. (2008) reports, that all organisms apart form plants and Apicomplexa exhibit GBD-FH3 containing formins in their genomes. All screened representatives within the present study contained at least one formin including the GBD-FH3 motif apart from all diatoms and the Apicomplexa *T. gondii* (Tab. 15.4). Diatoms and the Apicomplexa are therefore in demand for a regulatory region within their sequences. One could speculate that the multiple DUF region in the diatom formins assume a regulatory function throughout the cell cycle, although *Toxoplasma gondii* is not containing any formin including a DUF domain.

Incorporating all results the domain architecture of the formins leads to the assumption, that they are in general weak actin bundlers, but tight capping proteins. Some posses the ability to anchor F-actin to membranes and some are regulating the nuclear actin. Diatom formins are very probably not accelerating plus-end dynamics as in most other organisms due to the missing FH1 domain and the lack of the G-actin sequestering protein profilin. A particularly interesting formin isoform is Thaps 24787, which can provide a link to the microtubule cytoskeleton by its kinesin motor domain.

Gelsolin Superfamily Proteins

All screened diatoms encode at least one severin gene each, characterized by three replicated gelsolin repeat motifs (Fig. 15.9, 15.22). *P. tricornutum* codes for two severin genes with an identity of under 50% on the amino acid level. None of the screened diatom genomes/proteomes revealed a

member of the Gelsolin Superfamily with more than three gelsolin repeats. All diatom severin genes are likely to be fully functional as severing, capping and nucleating proteins.

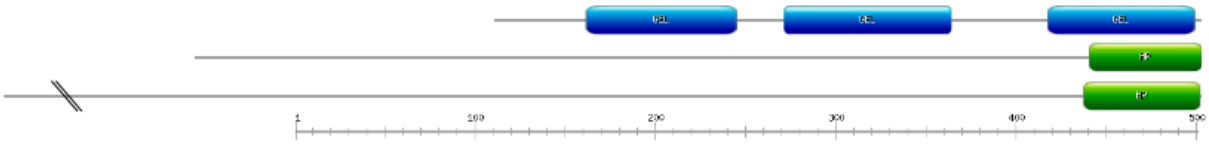


Figure 15.22: Illustrated secondary structure of the diatom severin at the top. All diatom severins are highly conserve therefore the secondary structure is only shown for one representative Fracy 271285. The middle shows the short dematin-like representative (Psemu 324055) and the bottom the long dematin-like (Phatr 45476) representative. The sequence of Psemu 324055 has been shortened by 1000 aa for illustration indicated by the double dash. Green — headpiece domain, blue — gelsolin domain.

Dematin

A C-terminal villin headpiece (HP) domain was identified within one *P. tricornutum* gene, one *P. multiseries* gene and two *F. cylindrus* genes (Fig. 15.23). The region upstream from the HP domain contained some elements of the dematin signature (Chen et al., 2009) such as the negatively charged region. However other typical dematin features like the increased number of prolines, or the PEST domain could not be confirmed within the dematin-like proteins of diatoms.

```
Phatr_45476       NREKYLSLSDFQAIF-GCNKDEFASWPKWKQTNLKRKAKLF- 559
DEMA_HUMAN_Q08495 RLERHLSAEDFSRVF-AMSPEEFGKLALWKRNELKKKASLF- 405
Psemu_324055      KWEEFLLEDEFVKHF-GVSKDEFYQQPKWKRDTLRRKIRVAF 1655
Fracy_244184      RWEYYLHDEEFQKCFNGLTKEQFYLQPKWKQDKQKWKLRVAF 1598
Fracy_241013      TWESYLNDEEFHKHFNGLTKDEFYLQPKWKRNRQKRNVRKAF 662
                    * .*  .:*   * . . ::*     **:   : :
```

Figure 15.23: Alignment of the C-terminal headpiece sequences of the four diatom dematin-like proteins with a human dematin protein. The actin-binding residues are highlighted in violet.

The two diatom dematin-like proteins Fracy 244184 and Psemu 324055 were relatively long for a dematin with around 1600 aa and the two smaller onces Phaeo 45476 and Fracy 241013 (about 600 aa) (Fig. 15.22) were still over 200 aa lager than the average dematin. The diatom dematin-like proteins are very likely to bind to actin as they contain a conserved HP domain (Fig. 15.23). Screening the four protein sequences for additional domains resulted only in one NLS (Fracy 244184), which makes this dematin-like protein a candidate for nuclear function. Dematins are membrane-cytoskeleton-associated proteins with F-actin-binding activity that induce F-actin bundle formation and stabilization, first described in (Siegel and Branton, 1985) and since then extensively studied in erythrocytes (Rana et al., 1993; Koshino et al., 2012; Wieschhaus et al., 2012; Brandt and Bailey, 2013). However, relatively little is known about the biological function of the dematins in non-erythroid cells (reviewed in Mohseni and Chishti, 2009). Therefore a function within the diatoms can not be predicted. Due to their up to four times larger size than an average dematin, it is quite possible that diatom dematin-like proteins assume a different function, than dematins in non-erythroid cells.

Notably only the pennate diatoms of the screened genomes encode a dematin-like gene, this might correlate with the stronger demand in pennate diatoms to bundle F-actin, because only in the pennates these thick actin bundles parallel to the raphe are found.

15.3.4 Bundlers & Crosslinkers

Apart from dematin, screening the diatom genomes/proteomes for additional actin bundlers gave hardly any results (Fig. 15.9). Blast searches for fascin genes resulted in hits for *F. cylindrus*, *T. pseudonana* and *T. oceanica*. Analyzing these BLAST hits for potential domains resulted by InterPro with a prediction of two actin cross-linking β- trefoil domains for the *F. cylindrus* and *T. oceanica* proteins and one β- trefoil domain for *T. pseudonana*. However, this prediction was not supported by MotifScann. Although the actin cross-linking β- trefoil domain might occur within these proteins of three of the diatom species, they are not fascin-like proteins. Fascin has four repeats of the β- trefoil domain whereas the diatom onces contained only one or two of these repeats. It might be possible that the proteins show some actin cross-linking activities, but this is also unlikely as the presence of these domains could be only confirmed by InterPro, but not be verified by MotifScann. Additional these genes have been classified as Interferon α-inducible protein 6/27, which are mitochondrial proteins and are therefore not likely to interact with the cytoplasmic F- actin.

That leaves only a very small set of actin-bundling proteins for diatoms with the dematin-like proteins of unknown function only detected in pennate diatom genomes and the formins. Diatoms have extensive F-actin bundling, such as the actin cables parallel to the raphe in pennates (chapter 7.1), and the actin ring at the leading edge of the SDV (Poulsen et al., 1999; Van de Meene and Pickett-Heaps, 2002; Tesson and Hildebrand, 2010a). It is, of course not entirely excluded, that diatoms possess additional actin-bundling proteins which have not been identified in the current screen. Though the chances for that are low, because the search criteria were based on the known structural features of actin binding proteins. It could be speculated, that diatoms posses unknown actin-bundling proteins which have been not described so far.

16 Conclusion

Gene duplication of ARPs and ABPs only occurred within *F. cylindrus*. The multiple gene copies of the gene family coded all for the same protein with a 100 % identity on the amino acid level. Mock (2013) previously reported a high number of gene duplication in various gene families throughout the genome of *F. cylindrus*. Among others this observation lead T. Mock to the speculation that *F. cylindrus* has lost the ability of sexual reproduction leading to a functional divergence of alleles. It is therefore likely, that the observed gene duplication of ARPs and ABPs are not the result of a specialization of the *F. cylindrus* actin cytoskeleton and its regulation, but rather a not yet understood gene duplication event occurring in *F. cylindrus*.

Considering the diversity and structural complexity of the diatom actin cytoskeleton (chapter 7.1), it was unexpected that the current screen revealed such a small set of ARPs and ABPs. For a dynamic and flexible actin cytoskeleton the depolymerization of F-actin, followed by recycling of ADP-actin to ATP-actin is the most important step. The diatom ABPs including ADF/cofilin, Aip1, coronin, severin (Fig. 16.1) and myosins (Heintzelman and Enriquez, 2010) should be sufficient for a fast disassembly of actin filaments (Blanchoin et al., 2014). Severed filaments could either serve as a new template for filament growth, or they could be further depolymerized. Up to this step the diatoms follow a well described path of recycling, although some organisms do employ a greater number of severing proteins. Where diatoms seem to have evolved a new mechanisms of regulation, is with regards to the control of the ATP-actin pool and the filament plus-end assembly. From the presented results it can be hypothesized, that diatoms developed a profilin-independent G-actin-binding and tethering mechanism. This assumption is not only supported by the lack of the conserved profilin binding site (P1) of most pennate CAP proteins, by the lack of the profilin-actin binding domain (FH1) of all of the many diatom formins, but furthermore by the absence of profilin (Fig. 15.13). The disassembled ADF/cofilin-bound ADP-actin monomer and ADP-actin is recycled to ATP-actin in diatoms very probably by CAPs. It can be therefore assumed that the ATP-actin pool is controlled by the CAPs. How the activated actin is dissociating from CAPs to be able to polymerize at the plus-end of the filament is unclear (Fig. 16.1). Also rapid elongation at the filament plus-end is probably not catalyzed through formins in diatoms (see 15.3.3) as there is no obvious domain (FH1) present in the diatom formins capable of recruiting G-actin. This indicates that ATP-actin polymerizes at the actin filament plus-end without the aid of known ABPs.

Assuming that the FH1-FH2-architecture represents the ancestral type of formin, the non-FH1 type may be considered as a derived or exotic variant of formin. Since the “exotic“ non-FH1 architecture prevails in Stramenopiles, but the presence or absence of profilin does not strictly correlate with that feature (Fig. 15.13), it follows that profilin still has a function relevant for the fitness of the organisms. The finding that also the P1 region is lost from the CAP proteins in six of the twelve screened Stramenopiles is in agreement with the presented general tendency to replace profilin-dependent mechanisms in the Stramenopiles. However, this evolutionary process

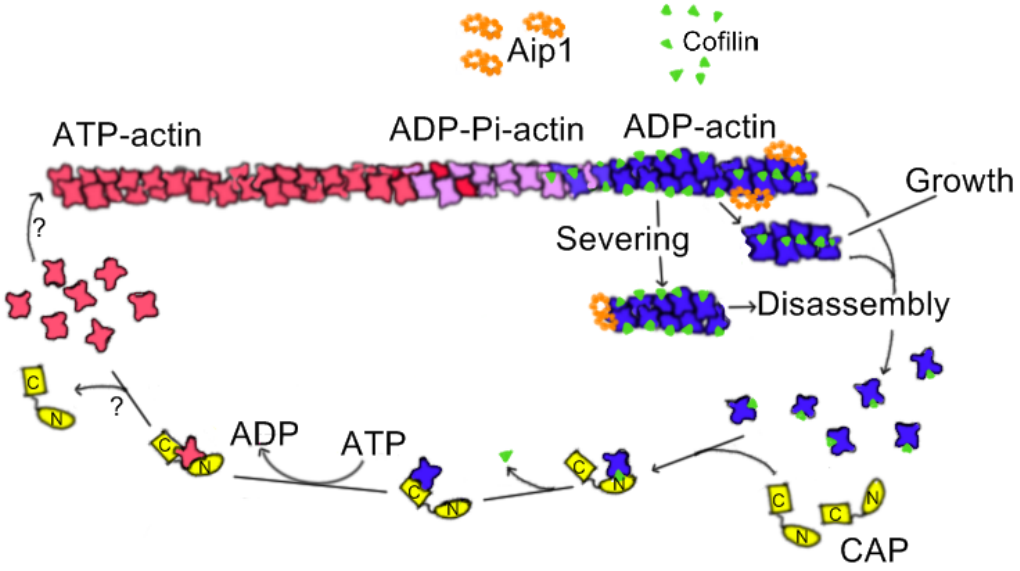


Figure 16.1: Schematic sketch of possible actin filament turnover in diatoms.

apparently has not come to an end, i.e. profilin function is not yet completely dispensable. This question could be further analyzed by correlating the occurrence of profilin with polyproline motifs in the proteins of the Stramenopiles.

Annexins are widely spread through the eukaryotic phyla. Therefore the differential distribution throughout the diatom genomes was an unexpected result. The centric diatoms alongside with the pennate species *P. tricornutum* possess two annexin genes, whereas the other pennates do not. It is possible that annexin functions are not as essential for diatoms as they are for other eukaryotes. Due to their unique silicified cell wall the cortical actin near the plasma membrane is quite likely not as dynamic as in other organisms (chapter 8). Even though diatoms do not have a very dynamic cortical actin, the actin cytoskeleton is most likely connected to the plasma membranes and further connected with cellular membranes throughout the cell cycle, i.e., to the SDV (chapter 5). Apart from formins, the annexins are the only other proteins that could be detected within the diatoms that are able to anchor actin to membranes (Fig. 15.9). Given such a general role it is notable that genomes of the two centrics encode annexin proteins and pennates do not, except for *P. tricornutum*. This diatom seems to hold a special position within the pennate diatoms.

A notably observation within the presented study was, that the presence and absence, the secondary structure, and the domain organization of *P. tricornutum* ARPs and ABPs, quite often seems to be more similar to the two centric diatoms, than to the other pennates. Then again even more striking is, that if one compares the sequence identity between the proteins, *P. tricornutum* clearly clusters with the pennate diatoms. This leads to the conclusion that evolutionary, the sequences of *P. tricornutum* genes clearly cluster with pennate diatoms, but that the regulation of the actin cytoskeleton is more similar to the centric diatoms. This observation can be only seen with regard to its unique, pleomorphic morphology (oval, fusiform or triradiate), which develops depending on the ambient environmental conditions (Tesson et al., 2009). The polymorphism of *P. tricornutum* probably lead to the need to maintain a domain structure in the ABPs and ARPs more alike to the centrics than to the pennates. It will be interesting to see, whether i.e. the two actin isoforms, or the two very similar formin isoforms are differentially expressed between the three morphotypes.

Summary

The cytoskeleton is essential for many cellular functions such as cell motility, the control of cell shape and polarity, meiosis, cytokinesis, intracellular transport as well as endo- and exocytosis. However, little is known about the structural organization and regulation of the microtubule (MT) and actin cytoskeleton in diatoms. This study presents for the first time immunolabeling of actin, α-tubulin, γ-tubulin as single label or in combinations, in a member of the class Bacillariophyceae, *Craticula cuspidata*.

The actin cytoskeleton of this species may be divided into three parts, a radially arranged fine cortical actin meshwork, a dens dynamic actin network in the deeper cytoplasm without obvious orientation and two thick prominent bundles parallel to the raphe. The actin drug, jasplakinolide allowed to gain insight into the dynamics of these structures. The results combined with previous publications indicate that the raphe bundles are highly dynamic structures recovering at a very fast speed after inhibitor treatment. Indicating, that actin filaments contributing to these bundles nucleate all along the raphe simultaneously rather than elongate unidirectional from the cell poles.

Co-localization of γ-tubulin and MTs demonstrated the existence of multiple MTOCs on the nuclear surface. MTs radiate out form the MTOCs into the periphery and further towards the cell poles, thus forming a highly symmetrical MT system. Oryzalin, a MT-depolymerizing drug, has a less severe effect on the MT system in *C. cuspidata*, than it is known from higher plants. It leads to a gradual depolymerization of MTs from the cell poles on the one hand, but induces the formation of MT aggregations particularly at the cell poles, on the other hand. These aggregates are located just next to the plastid tips in *C. cuspidata*, and as shown by electron microscopy correlate with a fibro-granular cup-like structure closely associated with the plastidal envelope of untreated cells. Although the affinity of these structures to microtubules suggests itself, γ-tubulin is not present in these structures and therefore cannot be responsible for that affinity.

Transformation of *C. cuspidata* with various GFP-fusion constructs failed for unknown reasons, however, the pennate species, *Cylindrotheca fusiformis*, was successfully transformed with the actin reporter, Lifeact-GFP. Expression of this construct allowed the visualization of all actin structure that were also labeled by immunofluorescence, such as the fine, cortical meshwork, the deeper cytoplasmic network and the thick bundles underneath the raphe. Now the experimental tool is available to address questions concerning the dynamics of the actin cytoskeleton and the connection of the actin cytoskeleton to the locomotion process of pennate diatoms.

Research on diatoms has advanced significantly over the last decade due to publicly available genome databases, yet up to now proteins involved in the regulation of the cytoskeleton in diatoms are largely unknown. Consequently, this study focuses on actin, actin-related proteins (ARPs) and actin-binding proteins (ABPs) encoded in the genomes of *Thalassiosira pseudonana*, *Thalassiosira oceanica*, *Phaeodactylum tricornutum*, *Fragilariopsis cylindrus* and *Pseudo-nitzschia multiseries*. The comparative genomic and phylogenetic study revealed, that most diatoms possess only a single conventional actin and a small set of ARPs. Among these are the highly conserved cytoplasmic Arp1 protein and the nuclear Arp4 as well as Arp6. The genomes of the diatoms contain two structurally different homologues of Arp4 that might serve specific functions. All diatom species, examined here, lack a homologue of the Arp3 protein and except for two species they also lack Arp2. Furthermore, no genes coding for other known protein components of the Arp2/3-complex could be found within the diatom genomes. These findings suggest that diatoms are not capable of forming the Arp2/3 complex, which is essential in most eukaryotes for actin filament branching and plus-end dynamics.

Diatoms encode a small set of ABPs, which should be efficient enough to regulate the disassembly of F-actin, the recycling of G-actin, as well as the capping of filaments and their anchoring to membranes. The secondary, and tertiary structures of the diatom ADF/cofilin, Aip1, and coronin proteins are all well conserved. Therefore, one can assume that disassembly through ADF/cofilin together with its enhancing and regulating partners Aip1 and coronin are fully functional within diatoms. The hypothesis is put forward that disassembled ADP-actin is recycled back to ATP-actin by CAPs, though it remains unclear, how the activated actin is dissociated from CAPs to become available for polymerization at the filament plus-ends.

Diatoms possess several multidomain variants of formin. All of them lack the FH1 domain suggesting that they are probably not capable of accelerating plus-end dynamics, a well studied function of formins in most other organisms. This feature coincides with the lack of the G-actin sequestering protein profilin, which would be the prime target of the FH1 domain. It therefore appears that diatoms have developed a novel, yet unknown way of filament growth and regulation of rapid filament elongation. This characteristic seems to have spread among the Stramenopiles, as most Stramenopila predominantly code for formins without a FH1 domain, but in diatoms it is most distinctly expressed.

Appendix

17 Chemicals & Enzymes

All chemicals and reagents used for the experiments were purchased from Bio-Rad Laboratories (Munich, Germany), Boehringer-Mannheim (Mannheim, Germany), Invitrogen (Darmstadt, Germany), Macherey Nagel GmbH (Dueren, Germany), Fermentas (New York, USA), Merck (Darmstadt, Germany), New England Biolabs (Frankfurt a. M., Germany), Pharmacia (Uppsala, S), Roth (Karlsruhe, Germany), Serva (Heidelberg, Germany), Sigma (Munich, Germany), SÜD-LABORBEDARF GmbH (Gauting, Germany), Thermo Fisher Scientific Inc., or Tropic Marin GmbH (Wartenberg, Germany) unless otherwise stated.

Table 17.1: PCR components and enzymes used in this study.

Enzyme	**Manufacturer**
Polymerase	
iProof	Bio-Rad Laboratories GmbH, Munich, Germany
Polymerase (Taq)	Omni Biolabs, Germany
dNTP, Phusion® High-Fidelity DNA Polymerase	Thermo Fisher Scientific Inc.
Restriction Enzymes and buffers	
Asc1, Bcl1, Bgl1, BspH1, BspT1, Dpn1, Hinc1, Oli1	Fermentas, Thermo Fisher Scientific Inc.
ApaI	Gibco®, Karlsruhe, Germany
AgeI	Promega GmbH, Mannheim, Germany
AccI, AleI, AscI, AvaII, BamHI, BlpI, BstEII, ClaI, ClaI Ro, DraI, EcoRI, HindIII, KpnI, NocI, NotI, PstI, SacI, SalI, SfiI, SmaI, SpeI, SphI, XbaI, XhoI, XmaI, XmnI	New England Biolabs, Frankfurt a. M., Germany

18 Appendix — Imunofluorescence labeling

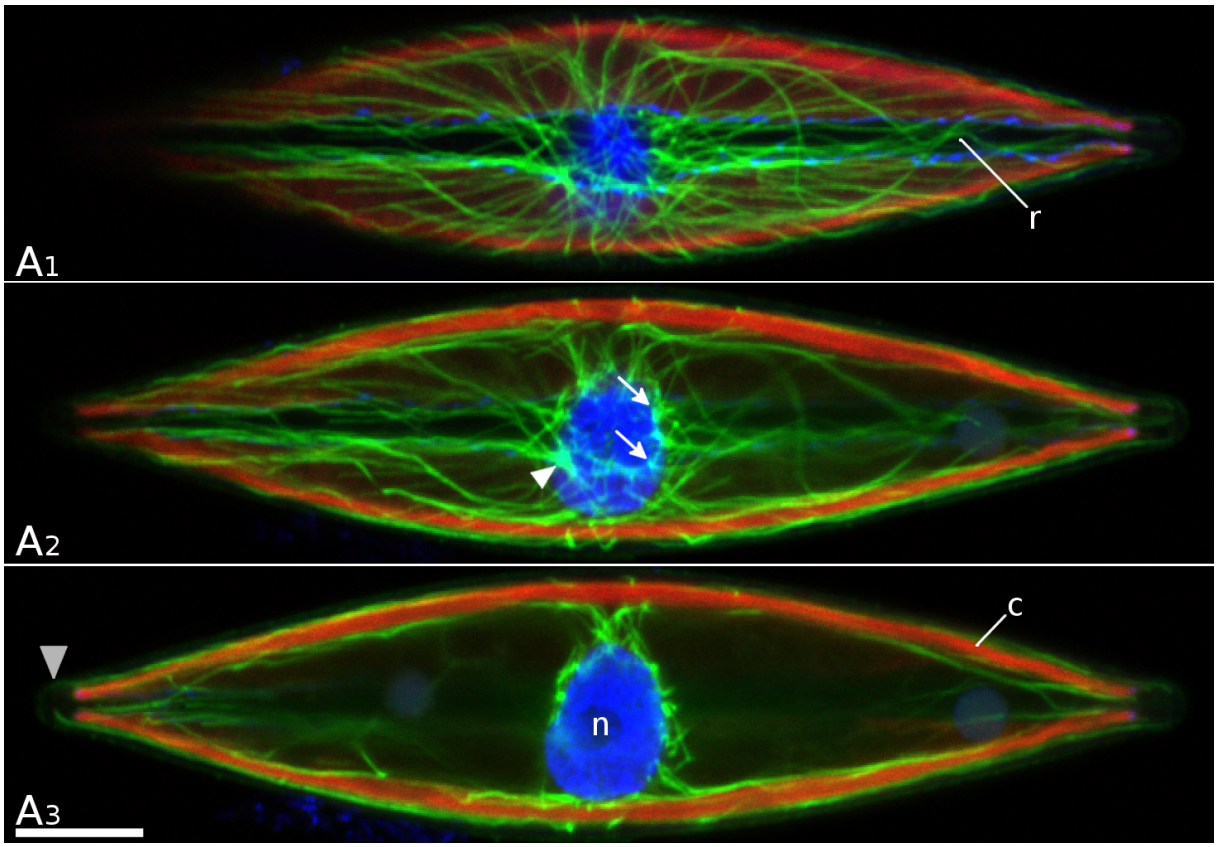


Figure 18.1: Labeling of α-tubulin in *C. cuspidata*. Single optical sections of the cell presented in Fig. 7.3 B as a projection. A1 — image about 2.1 μm under the plasma membrane, A2 — 0.7 μm inwards, A3 — image captured 1.4 μm further inwards. Symmetrically organized MTs radiating out from the MTOCs (arrow) and MCs (arrowheads) spanning the cell, but only slightly interspersed the cell poles (gray arrowhead), and spared the area under the raphe (r). c — plastid; n — nucleus; r — raphe; scale bar 10 μm.

19 Appendix — Transformation

19.1 Amplifying an endogenous actin promoter/terminator of *C. cuspidata*

19.1.1 Method

Amplifying an actin gene

The endogenous promoter and terminator of actin were chosen to construct a expression cassette to transform *C. cuspidata*. Two degenerated forward Primers AF1 and AF2 and three degenerated revers primers AR1, AR2 and AR3 were designed to amplify an actin gene of the unsequenced *C. cuspidata* genome. All possible primer combinations were used for the PCR reaction (Tab. 19.1). As a positive control, the actin genes of *A. thaliana* genome were amplify with same PCR program and components.

Table 19.1: PCR program and components used to amplify an actin gene with degenerated primers of the unsequenced C. cuspidata genome.

Components	Amount in μl	**Cycle - Temperature**	Time
DNA	1	1) 94℃	5 min
Buffer	5	2) 94℃	15 sec
Taq	0,5	3) 55℃	30 sec
dNTP	4	4) 72℃	3 min
Prim For	2	5) 72℃	10 min
Prim Rev	2	Repeat cycle 2-4 for 34 times	
MQ	35,5		

The resulting bands after gelelectrophoresis at a size between 500 bp and 1000 bp were excised from the gel and purified. The purified PCR products were diluted 1:100 with MQ and used as templates for a second PCR reaction with the same setting as above to amplify the sequences. The resulting bands of the second PCR were excised from the gel, purified and send to sequencing.

The sequenced DNA fragments were translated into a amino acid sequence with the Sequence Manipulation Suite: Translate (http://www.bioinformatics.org). The protein sequence was further verified to be actin by BLAST searches and Alignment (for details see 14).

Inverse PCR - Amplifying an endogenous actin promoter/terminator of *C. cuspidata*

The two successful obtained actin genes (*ActA* and *ActB*) were both chosen to amplify there promoter and terminator sequence of the unsequenced *C. cuspidata* genome. To amplify the up- and downstream region of these actin genes, the inverse PCR method (Ochman et al., 1988) was chosen. Therefore the genome of *C. cuspidata* was digested with various restriction enzymes, with one restriction enzyme per reaction. In total 36 different enzymes were tested (see 17.1). Restriction

Enzymes were used according to the manufacturer instructions and the digestion efficiency was controlled for each enzyme by gelelectrophoresis. After inhibiting the digestion, the genomic DNA-fragments were ligated over night at 16 °C and an inverse PCR followed, with the ligated genomic DNA-fragments as a template (Fig. 19.1).

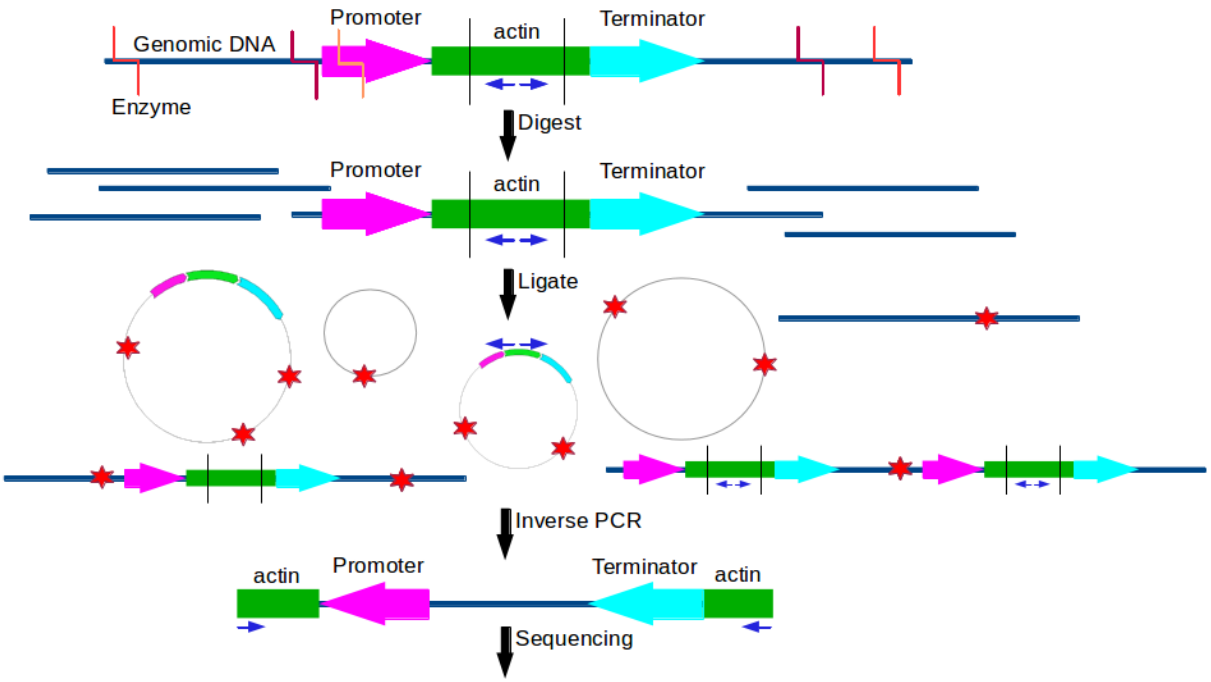


Figure 19.1: Schematic presentation of the Inverse PCR protocol. 36 different enzymes were used in 36 reactions to digest the genomic DNA at unknown positions. Ligation of the fragmented genomic DNA leads to a mixture of many different circulated fragments and ligated fragments; red stars - ligation site. A PCR product is only resulting when the gene of interest is circumcised to a vector, which is not exceeding the size of about 12 kbp.

The efficiency and specific of three different polymerases (iProof, Taq, Phusion®) were tested to improve the inverse PCR protocol with ligated DNA-fragments as a template. The concentrations of the components of the PCR reaction and the PCR program were modified and most parameters were changed (Tab. 19.2) to amplify the up- and downstream region of the coding actin gene.

As a negative control the digested genomic DNA without ligation, as well as genomic DNA were used as a template with the same PCR settings. For a positive control extracted genomic DNA of *A. thaliana* was digested with HindIII (cutting up- and downstream of the *Stonin* gene), followed by ligation (Tab. 19.3). Primers for an inverse PCR were designed for the Stonin gene *At2g20790* of *A. thaliana*. The same PCR settings as described in 19.2 were applied. To test the binding specific of the inverse PCR actin primers, only the forward or the revers primers were added to the PCR reaction.

Band pattens that reoccurred after two replicates of the same experimental settings and not appearing in the negative control were amplified by a second PCR (as described in 19.1.1), excised and send to sequencing.

Table 19.2: PCR program and component concentrations used to amplify the endogenous promoter/terminator of the unsequenced C. cuspidata genome. The minimum and the maximum of the modified spectrum of the concentrations, the temperatures and times that were attempted during this experiment are shown. When iProof, or Phusion® were used as a polymerase, DMSO was added to improve the PCR reaction.

Component	Amount in μl	**Cycle - Temperature**	Time
ligated DNA	0.1 - 2	1) 93 °C - 98 °C	30 sec - 5 min
Buffer	1/2	2) 93 °C - 98 °C	7 sec - 15 sec
Taq, iProof, Phusion®	0.1- 0.2	3) 48 °C - 57 °C	25 sec - 30 sec
dNTP	0.2 - 0.8	4) 72 °C	1 min - 6 min
Prim For	0.4 - 0.75	5) 72 °C	10 min
Prim Rev	0.4 - 0.75	Repeat cycle 2-4 for 30-34 times	
MQ	add to a volume of 10		

Table 19.3: Positive control of the Inverse PCR method, with the A. thaliana genome. Inverse PCR primers were designed for the *Stonin* gene. The genomic DNA was digested by HindIII, cutting several hundred base pairs up and downstream of the primer binding site.

Digestion	Amount in μl	**Ligation**	Amount in μl
DNA	8	digested DNA	15
Buffer	2.5	Buffer	5
HindIII	3	T4	1
ddH_2O	11.5	ddH_2O	29
for 3 h at 37 °C		over night at 16 °C	

19.1.2 Results

C. cuspidata actin gene

Two genomic DNA fragments were successfully amplified with the degenerate primers designed for actin. The primer pair AF2-AR2 and AF1-AR3 both resulted in a single band PCR product. The two products are slight differences in size with one around 700 bp and the other at about 550 bp (Fig. 19.2). The sequence of the two amplified genes confirmed them to be fragments of two different actin genes. The same two primer pairs amplified actin genes of the *A. thaliana* genome, that are due to an intron about 200 bp lager.

C. cuspidata actin gen A (*ActA*) was amplified with the primer pair AF2-AR2. Although the amplified gene part was about 550 bp the information of only 497 bp is available, due to sequencing arrows up- and downstream of the core part (Tab. 19.4). Corresponding to that a 602 bp long core sequence was available of the about 700 bp long actin gen B (*ActB*), amplified with the primer pair AF1-AR3.

Inverse PCR - Amplifying an endogenous actin promoter/terminator of *C. cuspidata*

The approach to amplify the endogenous promoter and terminator of *ActA* and *ActB* by inverse PCR failed. To recognize the inverse PCR product as the amplified up- and downstream region of the actin gene, the inverse PCR primers were chosen to be in the center of the actin gene. So at least 100 bp of the up- and downstream region of *ActA* or *ActB* should be part of the positive

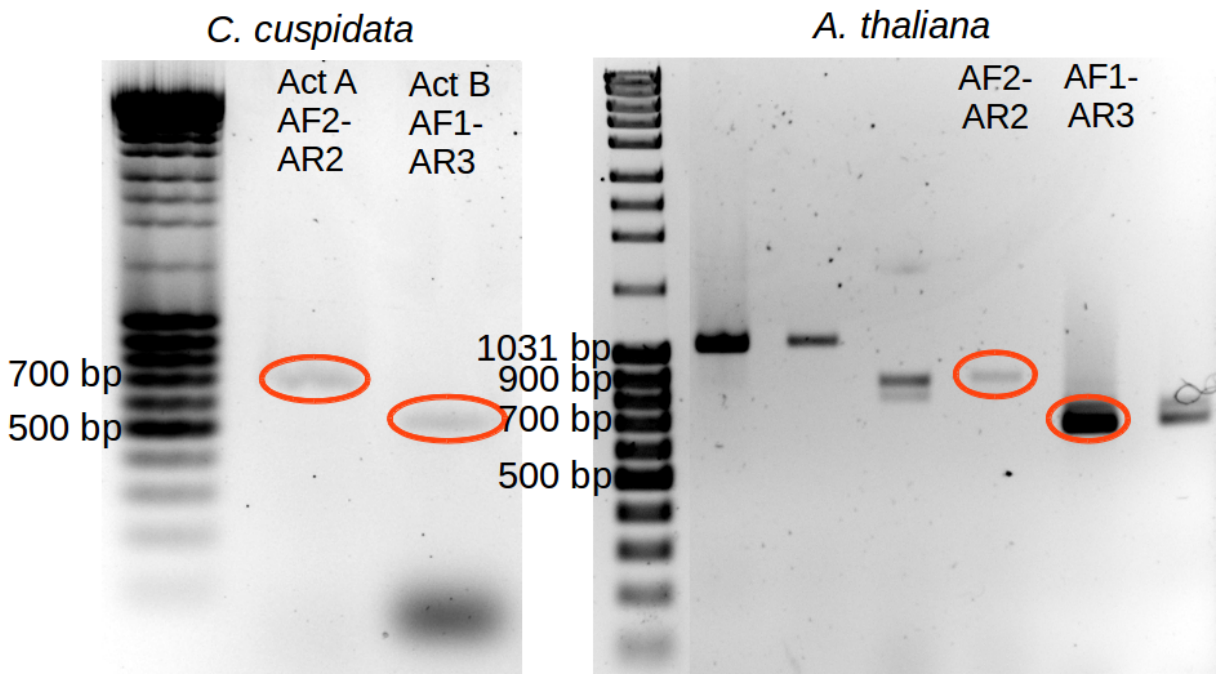


Figure 19.2: Results of the actin amplification of the *C. cuspidata* genome in comparison to the *A. thaliana* genome. The red encircled bands were send for sequencing.

Table 19.4: Sequence of the two C. cuspidata actin genes ActA and ActB.

ActA
CTATAGAGGCTCACCCTGTTCTCCTCACGGAAGCTCCCCAAAATCCTAAGGCTAACCGCG
AGCGCATGACACAAATCATGTTCGAGACTTTCAACGTCCCCGCCATGTATGTGAACATTC
AGGCAGTTTTGTCTCTGTACGCTTCGGGTCGTACTACTGGATGCGTGTTGGACTCTGGTG
ATGGTGTAACCCACACTGTGCCGATTTACGAAGGTTACGCCCTTCCTCATGCCGTCATTC
GTTTGGATCTTGCTGGACGTGACTTGACCGATTACTTGATGAAGATTCTCACGGAACGTG
GTTACTCTTTGACCACCACGGCAGAGCGTGAAATTGTTCGAGATATCAAGGAATCGCTTT
GCTTCGTTGCAGTAGACTTTGAGGAGGAAATGAAAAAGGCAGCTGAATCTTCGGCTCTTG
AGAAGTCTTTCGAACTTCCCGATGGTAACATCATTGTGATTGGAAACGAACGCTTCCGTT GCC-
CGGAGGTCTCTTTTC
ActB
CAGTTTTGTTGACCGAAGCCCCAATGAACCCTAAATCTAACCGTGAAAAGATGACCCAAA
TCATGTTTGAAACCTTCAACGTACCAGCCTTCTACGTTTCTATCCAAGCCGTCTTATCTT TAT-
ACTCGTCTGGTAGAACCACCGGTATCGTTTTAGATTCGGGTGATGGTGTTACTCACG TTGTTC-
CAATTTACGCCGGTTTCTCCTTACCACACGGTATCTTAAGAATTGACTTGGCTG GTAGA-
GACTTGACCGACTATTTGATGAAGATCTTGTCCGAAAGAGGTTACACTTTCTCCA CCACCGCT-
GAAAGAGAAATTGTCCGTGATATCAAAGAAAAATTATGTTACGTTGCCTTGG ACTTTGAACAA-
GAAATGCAAACTTCATCTCAATCCTCCGCCATCGAAAAGTCTTACGAAT TACCTGATGGCCAAGT-
TATTACTATCGGTAACGAAAGATTTAGAGCTTCCGAAGCTTTGT TCCGTCCTTCTGACTTAGGTTTA-
GAAGCCGCTGGTATTGACCAAACCACTTACAACTCGA TTATGAAGTGTGATGTCGATGTCAGAAAG-
GAATTATACGGTAACATTGTTATGTCTGGTG GTACCCCATT

inverse PCR product and therefore easily recognizable. None of the sequenced inverse PCR products aligned partial to the *ActA* or *ActB* sequence.

The genome of *C. cuspidata* was digested by 36 different restriction enzymes, meaning 36 different samples resulted as a template for the inverse PCR. The inverse PCR was repeated for each sample at least three times, only few samples resulted in a repeatable band patterning, that

was not shown in the negative controls. These bands were excised and send to sequencing, but could not be verified as actin containing sequences. The negative control (digested *C. cuspidata* DNA without ligation) resulted in a strong band patterning, most of them identical to the results of the samples (Fig. 19.3). Testing the primers towards there specifics, with genomic DNA as a template resulted in a strong band patten, even when only the forward, or revers primer was added to the reaction (Fig. 19.5). Sequencing the band at 1000 bp amplified by ActAIF1 (Fig. 19.5) did not result in any actin gene containing sequence.

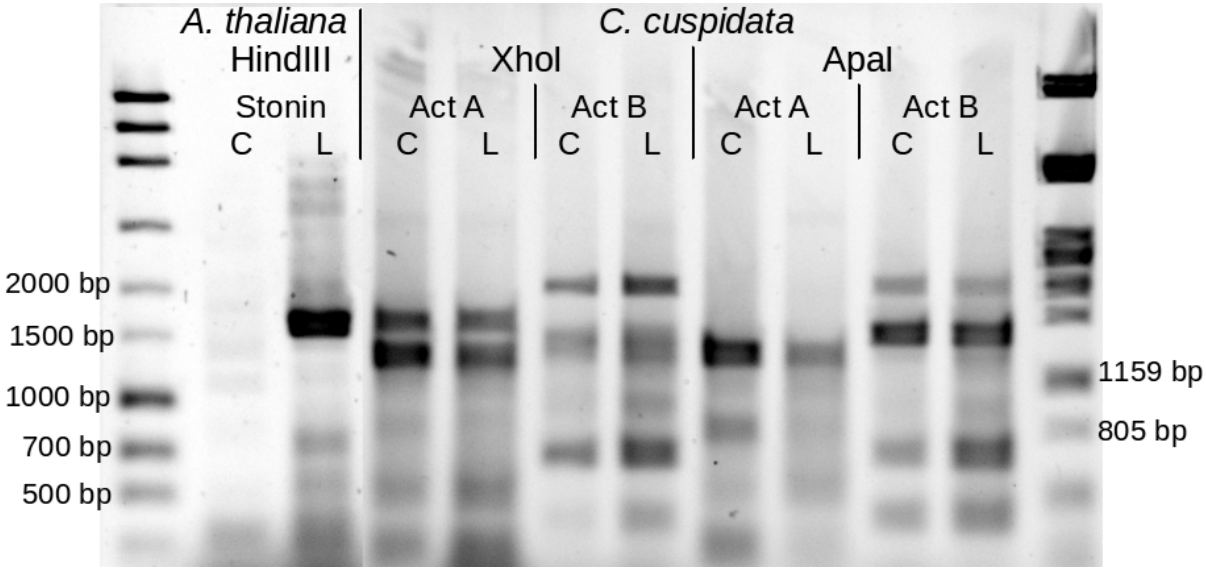


Figure 19.3: Positive and negative control of the inverse PCR. The positive control with the *A. thaliana* stonin gene resulted in a clear thick band at the expected size of 1572 bp. The negative control with digested *C. cuspidata* DNA resulted in the same patterning as the digested and ligated DNA samples. Used restriction enzymes shown here are HindIII, XhoI and ApaI. C — digested DNA (control), L — digested and ligated DNA.

The positive control of the inverse PCR, with primers for the *Stonin* of *A. thaliana* and genomic DNA of *A. thaliana* as a template, resulted in a strong single band, matching the predicted size of 1572 bp (Fig. 19.3).

The different polymerases Taq, iProof and Phusion resulted in different band pattern, where Taq is amplifying the fewest bands and Phusion the most. Both proofreading polymerases iProof and Phusion showed a similar patten (Fig. 19.6). The band patterning was further compared with the corresponding negative control. Bands that were not amplified in the negative control have been send to sequencing. None of the three sequences corresponded to an actin sequence.

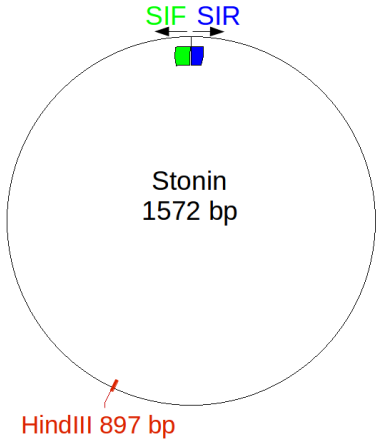


Figure 19.4: Inverse PCR - positive control. The resulting vector (1572 bp) of the *Stonin* gene after digestion and ligation with the two Inverse PCR primers SIF-SIR.

After various approaches testing different parameters and settings the attempt to amplify an actin promoter and terminator of *C. cuspidata* was resigned.

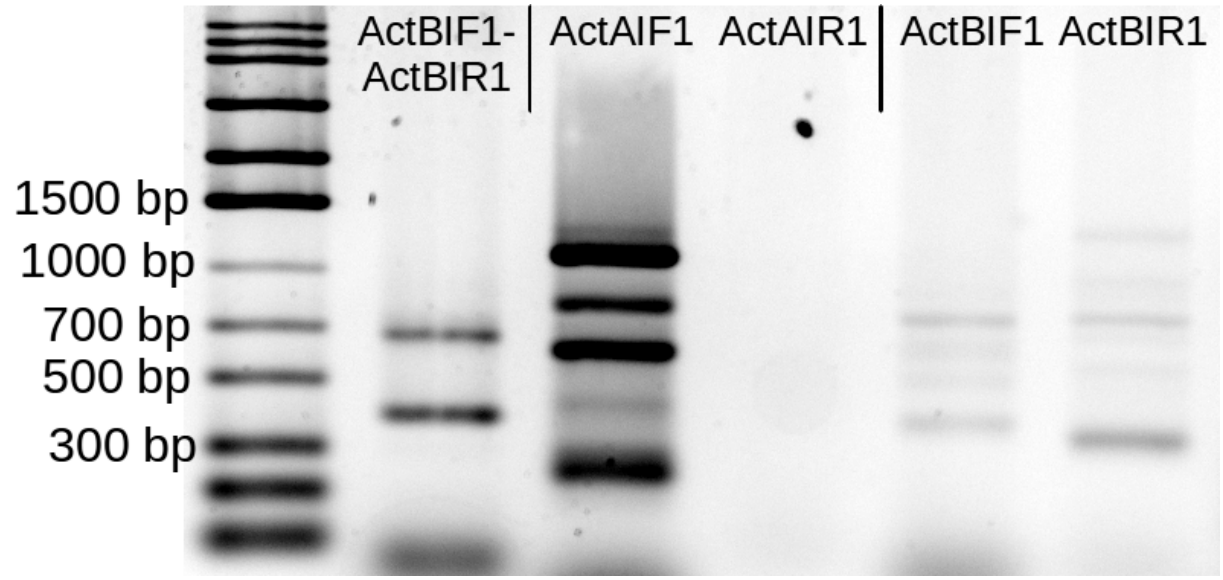


Figure 19.5: Binding and amplification ability of the inverse PCR primers with genomic *C. cuspidata* DNA. Adding only the forward primer ActAIF1 to the PCR reaction resulted in strong band pattens. The band at 1000 bp was excised and sequenced. The revers primer ActAIR1 was the only primer that did not amplify a sequence from the genomic DNA.

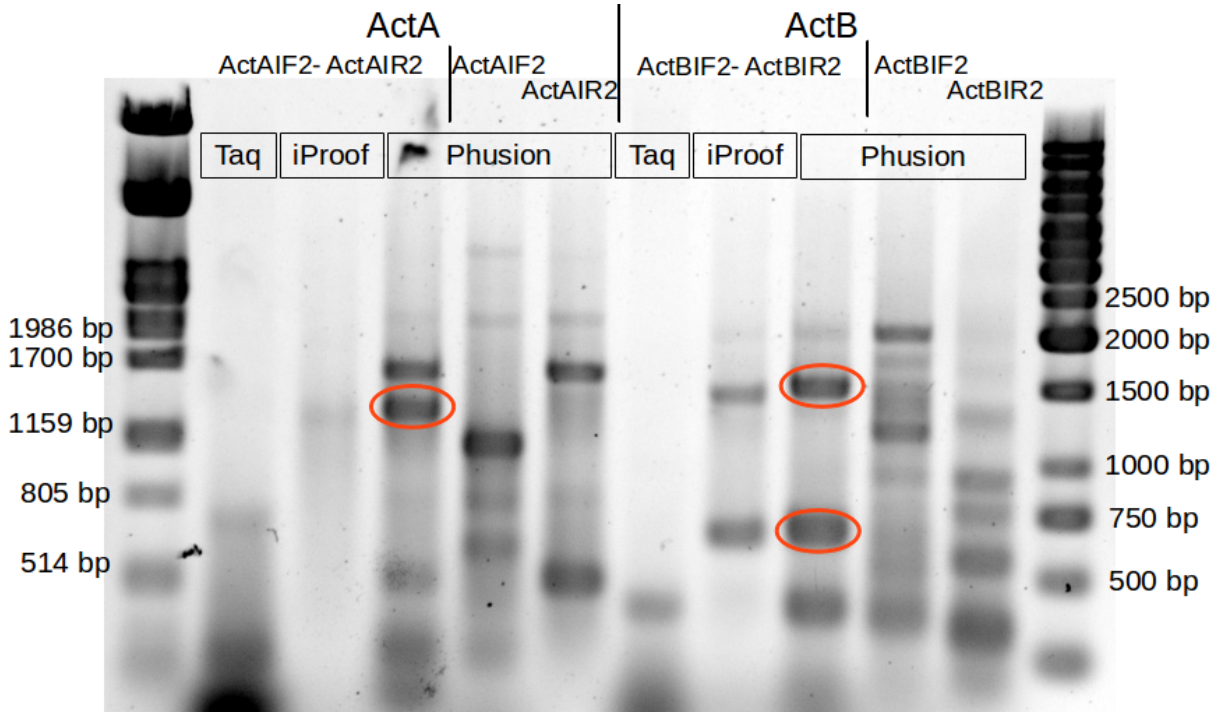


Figure 19.6: Comparison of the amplification ability of the three polymerases Taq, iProof and Phusion® resulted in different band patterning, although the same PCR components and settings were used for all reactions. The template was genomic DNA digested with AccI followed by ligation. The tree red encircled bands were sent for sequencing.

20 Appendix — Comparative genomics

Table 20.1: List of proteins with JGI Protein-ID. Gene models that have been manual modified are indicated. Gene models with and different Protein-ID, presented the same gene are listed as identical gene models. GM — gene model

Protein	Species	ID	Allelic variance	Gene Duplication	Modified GM	Identical GM
Actin	Fracy	171145	176600	269912,209894,176600, 222567,214941		276129
	Psemu	22272		216762		
	Phaeo	29136				
		51157		29812		
	Thaps	25772				
	Thaoc	16346				
Arp1	Fracy	228346				247121
	Psemu	239296				
	Phaeo	44089				
	Thaps	269504			yes	
	Thaoc	31054				
Arp2	Fracy	269231	254983			187716
	Psemu	288992				
Arp4	Fracy	178989			yes	
		188056			yes	188033
		198035			yes	
	Psemu	188258				
		324603				
	Phaeo	20837				
		4515				
	Thaps	269619			yes	
		41068				
	Thaoc	22655			Fragment	

		11934			Fragment	
Arp6	Fracy	139659			yes	
	Psemu	62991				
	Phaeo	3281			yes	
	Thaps	260856			yes	
	Thaoc	3168			Fragment	
Arp11	Thaps	1132				
	Thaoc	19871				
ADF/cofilin	Fracy	206967				
	Psemu	248251		248057		
	Phaeo	15613			yes	
	Thaps	4830				
	Thaoc	9491				
Aip1	Fracy	185021	199611			22581
	Psemu	260507				
	Phaeo	48381				
	Thaps	23728				7517
	Thaoc	5008				
Annexin	Phaeo	44109				33469
		54190				
	Thaps	259626				
		268316				
	Thaoc	10783				
		19272			Fragment	
CAP	Fracy	252941				
		262057	252941			
	Psemu	249042			yes	
	Phaeo	20921				13047
	Thaps	23046			yes	34992
	Thaoc	939				
CAPZ	Fracy	234298				
		246923		Isoform 246175		

	Psemu	304923				
		36278				
	Phaeo	35252				
		9601			yes	
	Thaoc	14480			Fragment	
		34942			Fragment	
Coronin	Fracy	190898		189497		
	Psemu	70315				
	Phaeo	45580				
	Thaps	8387				
	Thaoc	35123			yes	
Dematin	Fracy	244184	254255			
		241013	251628			
	Psemu	324055				
	Phaeo	45476				
Fascin	Fracy	235712	Fascin			
		254447				
	Thaps	7140				
	Thaoc	15375				
		5944				
Formin	Fracy	246977	193211			
		163042	153464		yes	
	Psemu	212264			yes	
		260762			yes	
	Phaeo	54510				
		46029		Isoform 54510		
		54518				
		54229				
		46058				
	Thaps	4119				
		9248				
		11711				
		24787			yes	9733

		8340	24141
		24070	8191; 263598
	Thaoc	K0R7S7	
		K0RS69	
		K0RCQ7	
		K0TL59	
		K0RD64	
		K0QZG3	
		K0R2S6	
Severin	Fracy	271285	
	Psemu	209419	
	Phaeo	53980	
		19089	
	Thaps	22984	
	Thaoc	21027	

Bibliography

Amin, S. A., Parker, M. S., and Armbrust, E. V. (2012). Interactions between diatoms and bacteria. *Microbiol. Mol. Biol. Rev.*, 76(3):667–84.

André, E., Lottspeich, F., Schleicher, M., and Noegel, a. (1988). Severin, gelsolin, and villin share a homologous sequence in regions presumed to contain F-actin severing domains. *J. Biol. Chem.*, 263(2):722–7.

Apt, K., Grossman, A., and Kroth-Pancic, P. (1996). Stable nuclear transformation of the diatom Phaeodactylum tricornutum. *Mol. Gen. Genet.*, 252:572–9.

Aravind, L. and Koonin, E. (2000). SAP–a putative DNA-binding motif involved in chromosomal organization. *Trends Biochem. Sci.*, 25(3):112–4.

Armbrust, E. V., Berges, J. a., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., Zhou, S., Allen, A. E., Apt, K. E., Bechner, M., Brzezinski, M. a., Chaal, B. K., Chiovitti, A., Davis, A. K., Demarest, M. S., Detter, J. C., Glavina, T., Goodstein, D., Hadi, M. Z., Hellsten, U., Hildebrand, M., Jenkins, B. D., Jurka, J., Kapitonov, V. V., Kröger, N., Lau, W. W. Y., Lane, T. W., Larimer, F. W., Lippmeier, J. C., Lucas, S., Medina, M., Montsant, A., Obornik, M., Parker, M. S., Palenik, B., Pazour, G. J., Richardson, P. M., Rynearson, T. a., Saito, M. a., Schwartz, D. C., Thamatrakoln, K., Valentin, K., Vardi, A., Wilkerson, F. P., and Rokhsar, D. S. (2004). The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism. *Science*, 306:79–86.

Aumeier, C. and Menzel, D. (2012). Secretion in the diatoms. *Secretions exudates Biol. Syst.*, 12:1–30.

Baarlink, C. and Grosse, R. (2014). Formin' actin in the nucleus. *Nucleus*, 5(1):1–6.

Baird, W. and Meagher, R. (1987). A complex gene superfamily encodes actin in petunia. *EMBO J.*, 6(11):3223–31.

Baker, K. H. and Herson, D. S. (1978). Interactions between the diatom Thalassiosira pseudonana and an associated Pseudomonad in a mariculture system. *Appl. Environ. Microbiol.*, 35(4):791–6.

Balcer, H. I., Goodman, A. L., Rodal, A. a., Smith, E., Kugler, J., Heuser, J. E., and Goode, B. L. (2003). Coordinated regulation of actin filament turnover by a high-molecular-weight Srv2/CAP complex, Cofilin, Profilin, and Aip1. *Curr. Biol.*, 13:2159–69.

Bamburg, J. R. (1999). Proteins of the ADF/COFILIN Family : Essential Regulators of Actin Dynamics. *Annu. Rev. Cell Dev. Biol.*, 15:185–230.

Bar-Nun, S., Shneyour, Y., and Beckmann, J. (1983). G-418, an elongation inhibitor of 80 S ribosomes. *Biochim Biophys Acta.*, 741(1):123–7.

Barton, G. J., Newmann, R. H., Freemont, P. S., and Crumpton, M. J. (1991). Amino acid sequence analysis of the annexin super gene family of proteins. *Eur. J. Biochem.*, 198(3):749–60.

Beltzner, C. C. and Pollard, T. D. (2008). Pathway of actin filament branch formation by Arp2/3 complex. *J. Biol. Chem.*, 283(11):7135–44.

Bertling, E., Quintero-Monzon, O., Mattila, P. K., Goode, B. L., and Lappalainen, P. (2007). Mechanism and biological role of profilin-Srv2/CAP interaction. *J. Cell Sci.*, 120(7):1225–34.

Bhattacharya, D., Aubry, J., Twait, E. C., and Jurk, S. (2000). Actin gene duplication and the evolution of morphological complexity in land plants. *J. Phycol.*, 36:813–20.

Blackbourn, H. D. and Battey, N. H. (1993). Annexin-mediated secretory vesicle aggregation in plants. *Physiol. Plant.*, 89:27–32.

Blanchoin, L., Amann, K. J., Higgs, H. N., Marchand, J. B., Kaiser, D. a., and Pollard, T. D. (2000). Direct observation of dendritic actin filament networks nucleated by Arp2/3 complex and WASP/Scar proteins. *Nature*, 404(27):1007–11.

Blanchoin, L., Boujemaa-Paterski, R., Henty, J. L., Khurana, P., and Staiger, C. J. (2010). Actin dynamics in plant cells: a team effort from multiple proteins orchestrates this very fast-paced game. *Curr. Opin. Plant Biol.*, 13:714–23.

Blanchoin, L., Boujemaa-Paterski, R., Sykes, C., and Plastino, J. (2014). Actin dynamics, architecture, and mechanics in cell motility. *Physiol. Rev.*, 94(1):235–63.

Blanchoin, L. and Pollard, T. (1999). Mechanism of interaction of Acanthamoeba actophorin (ADF/Cofilin) with actin filaments. *J. Biol. Chem.*, 274(22):15538–46.

Borisy, G. and Svitkina, T. (2000). Actin machinery: pushing the envelope. *Curr. Opin. Cell Biol.*, 12:104–11.

Bowler, C., Allen, A. E., Badger, J. H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otillar, R. P., Rayko, E., Salamov, A., Vandepoele, K., Beszteri, B., Gruber, A., Heijde, M., Katinka, M., Mock, T., Valentin, K., Verret, F., Berges, J. A., Brownlee, C., Cadoret, J. P., Chiovitti, A., Choi, C. J., Coesel, S., De Martino, A., Detter, J. C., Durkin, C., Falciatore, A., Fournet, J., Haruta, M., Huysman, M. J., Jenkins, B. D., Jiroutova, K., Jorgensen, R. E., Joubert, Y., Kaplan, A., Kroger, N., Kroth, P. G., La Roche, J., Lindquist, E., Lommer, M., Martin-Jezequel, V., Lopez, P. J., Lucas, S., Mangogna, M., McGinnis, K., Medlin, L. K., Montsant, A., Oudot-Le Secq, M. P., Napoli, C., Obornik, M., Parker, M. S., Petit, J. L., Porcel, B. M., Poulsen, N., Robison, M., Rychlewski, L., Rynearson, T. A., Schmutz, J., Shapiro, H., Siaut, M., Stanley, M., Sussman, M. R., Taylor, A. R., Vardi, A., von Dassow, P., Vyverman, W., Willis, A., Wyrwicz, L. S., Rokhsar, D. S., Weissenbach, J., Armbrust, E. V., Green, B. R., de Peer, Y., and Grigoriev, I. V. (2008). The Phaeodactylum genome reveals the evolutionary history of diatom genomes. *Nature*, 456:239–44.

Brandt, G. S. and Bailey, S. (2013). Dematin, a human erythrocyte cytoskeletal protein, is a substrate for a recombinant FIKK kinase from Plasmodium falciparum. *Mol. Biochem. Parasitol.*, 191(1):20–3.

Brown, R. and Lemmon, B. (2007). The pleiomorphic plant MTOC: an evolutionary perspective. *J. Integr. Plant Biol.*, 49(8):1142–53.

Bruckner, C. G., Bahulikar, R., Rahalkar, M., Schink, B., and Kroth, P. G. (2008). Bacteria associated with benthic diatoms from Lake Constance: phylogeny and influences on diatom growth and secretion of extracellular polymeric substances. *Appl. Environ. Microbiol.*, 74(24):7740–9.

Bubb, M. R. (2000). Effects of Jasplakinolide on the Kinetics of Actin Polymerization. An explanation for certain in vivo observations. *J. Biol. Chem.*, 275(7):5163–5170.

Bushell, E. S. C., Ecker, A., Schlegelmilch, T., Goulding, D., Dougan, G., Sinden, R. E., Christophides, G. K., Kafatos, F. C., and Vlachou, D. (2009). Paternal effect of the nuclear formin-like protein MISFIT on Plasmodium development in the mosquito vector. *PLoS Pathog.*, 5(8):e1000539.

Caldwell, J., Heiss, S., Mermall, V., and Cooper, J. (1989). Effects of CapZ, an Actin Capping Protein of Muscle, on the Polymerization of Actin. *Biochemistry*, 28(21):8506–14.

Carlier, M.-F., Ressad, F., and Pantaloni, D. (1999). Control of Actin Dynamics in Cell Motility: Role of ADF/COFILIN. *J. Biol. Chem.*, 274(48):33827–30.

Carroll, A., Moyen, C., Van Kesteren P, Tooke, F., Battey, N., and Brownlee, C. (1998). Ca2+, annexins, and GTP modulate exocytosis from maize root cap protoplasts. *Plant Cell*, 10:1267–76.

Casella, J. and Torres, M. (1994). Interaction of Cap Z with Actin. *J. Biol. Chem.*, 269(9):6992–8.

Chang, F., Drubin, D., and Nurse, P. (1997). Cdc12P, a Protein Required for Cytokinesis in Fission Yeast, Is a Component of the Cell Division Ring and Interacts With Profilin. *J. Cell Biol.*, 137(1):169–82.

Chaudhry, F., Breitsprecher, D., Little, K., Sharov, G., Sokolova, O., and Goode, B. L. (2013). Srv2/cyclase-associated protein forms hexameric shurikens that directly catalyze actin filament severing by cofilin. *Mol. Biol. Cell*, 24(1):31–41.

Chaudhry, F., Guérin, C., von Witsch, M., Blanchoin, L., and Staiger, C. J. (2007). Identification of Arabidopsis cyclase-associated protein 1 as the first nucleotide exchange factor for plant actin. *Mol. Biol. Cell*, 18:3002–14.

Chaudhry, F., Little, K., Talarico, L., Quintero-Monzon, O., and Goode, B. L. (2010). A central role for the WH2 domain of Srv2/CAP in recharging actin monomers to drive actin turnover in vitro and in vivo. *Cytoskeleton (Hoboken).*, 67(2):120–33.

Chen, L., Jiang, Z. G., Khan, A. a., Chishti, A. H., and McKnight, C. J. (2009). Dematin exhibits a natively unfolded core domain and an independently folded headpiece domain. *Protein Sci.*, 18(3):629–36.

Chepurnov, V. and Mann, D. (2004). Auxosporulation of Licmophora communis (Bacillariophyta) and a review of mating systems and sexual reproduction in araphid pennate diatoms. *Phycol. Res.*, 52:1–12.

Chesarone, M., DuPage, A., and Goode, B. (2009). Unleashing formins to remodel the actin and microtubule cytoskeletons. *Nat. Rev. Mol. cell Biol.*, (11):62–74.

Cohn, S., Nash, J., and Pickett-Heaps, J. (1989). The effect of drugs on diatom valve morphogenesis. *Protoplasma*, 149:130–43.

Condeelis, J. (2001). How is actin polymerization nucleated in vivo? *Trends Cell Biol.*, 11:288–293.

Cooley, L., Verheyen, E., and Ayers, K. (1992). chickadee encodes a profilin required for intercellular cytoplasm transport during Drosophila oogenesis. *Cell*, 69:173–84.

Cooper, J. (1987). Effects of cytochalasin and phalloidin on actin. *J. Cell Biol.*, 105:1473–8.

Cooper, J. and Pollard, T. (1985). Effect of capping protein on the kinetics of actin polymerization. *Biochemistry*, 24(3):793–9.

Cooper, J. a. and Schafer, D. a. (2000). Control of actin assembly and disassembly at filament ends. *Curr. Opin. Cell Biol.*, 12:97–103.

Coué, M., Brenner, S., Spector, I., and Korn, E. (1987). Inhibition of actin polymerization by latrunculin A. *FEBS Lett.*, 213(2):316–8.

Courtemanche, N. and Pollard, T. D. (2012). Determinants of Formin Homology 1 (FH1) domain function in actin filament elongation by formins. *J. Biol. Chem.*, 287:7812–20.

Crowell, E., Bischoff, V., Desprez, T., Rolland, A., Stierhof, Y.-D., Schumacher, K., Gonneau, M., Höfte, H., and Vernhettes, S. (2009). Pausing of Golgi bodies on microtubules regulates secretion of cellulose synthase complexes in Arabidopsis. *Plant Cell*, 21(4):1141–54.

Cvrcková, F., Novotný, M., Pícková, D., and Zárský, V. (2004). Formin homology 2 domains occur in multiple contexts in angiosperms. *BMC Genomics*, 5(1):44.

Cvrčková, F. (2013). Formins and membranes: anchoring cortical actin to the cell wall and beyond. *Front. Plant Sci.*, 4(November):436.

De Martino, A., Amato, A., and Bowler, C. (2009). Mitosis in diatoms: rediscovering an old model for cell division. *BioEssays*, 31:874–84.

DesMarais, V., Ghosh, M., Eddy, R., and Condeelis, J. (2005). Cofilin takes the lead. *J. Cell Sci.*, 118(1):19–26.

Dion, V., Shimada, K., and Gasser, S. M. (2010). Actin-related proteins in the nucleus: life beyond chromatin remodelers. *Curr. Opin. Cell Biol.*, 22:383–91.

Dodatko, T., Fedorov, A. a., Grynberg, M., Patskovsky, Y., Rozwarski, D. a., Jaroszewski, L., Aronoff-Spencer, E., Kondraskina, E., Irving, T., Godzik, A., and Almo, S. C. (2004). Crystal structure of the actin binding domain of the cyclase-associated protein. *Biochemistry*, 43(33):10628–41.

Dominguez, R. (2004). Actin-binding proteins–a unifying hypothesis. *Trends Biochem. Sci.*, 29(11):572–8.

Dominguez, R. (2010). Structural insights into de novo actin polymerization. *Curr. Opin. Struct. Biol.*, 20(2):217–25.

Dominguez, R. and Holmes, K. C. (2011). Actin structure and function. *Annu. Rev. Biophys.*, 40:169–86.

dos Remedios, C. G., Chhabra, D., Kekic, M., Dedova, I. V., Tsubakihara, M., Berry, D. a., and Nosworthy, N. J. (2003). Actin binding proteins: regulation of cytoskeletal microfilaments. *Physiol. Rev.*, 83:433–73.

Drocourt, D., Calmels, T., Reynes, J., Baron, M., and Tiraby, G. (1990). Cassettes of the Streptoalloteichus hindustanus ble gene for transformation of lower and higher eukaryotes to phleomycin resistance. *Nucleic Acids Res.*, 18(13):4009.

Drum, R. W. (1963). The cytoplasmic fine structure of the diatom, Nitzschia palem. *J. Cell Biol.*, 18:429–40.

Dunahay, T. G., Jarvis, E. E., and Roessler, P. G. (1995). Genetic transformation of the diatoms Cyclotella cryptica and Navicula Saprophila. *J. Phycol.*, 31:1004–12.

Eckley, D. M. and Schroer, T. A. (2003). Interactions between the Evolutionarily Conserved, Actin-related Protein, Arp11, Actin, and Arp1. *Mol. Biol. Cell*, 14(July):2645–54.

Edgar, L. A. and Pickett-Heaps, J. D. (1984). Diatom locomotion. *Prog. Phycol. Res.*, 3:47–88.

Eichinger, L., Noegel, a. a., and Schleicher, M. (1991). Domain structure in actin-binding proteins: expression and functional characterization of truncated severin. *J. Cell Biol.*, 112(4):665–76.

Era, A., Tominaga, M., Ebine, K., Awai, C., Saito, C., Ishizaki, K., Yamato, K. T., Kohchi, T., Nakano, A., and Ueda, T. (2009). Application of Lifeact reveals F-actin dynamics in Arabidopsis thaliana and the liverwort, Marchantia polymorpha. *Plant Cell Physiol.*, 50(6):1041–8.

Esue, O., Harris, E. S., Higgs, H. N., and Wirtz, D. (2008). The filamentous actin cross-linking/bundling activity of mammalian formins. *J. Mol. Biol.*, 384(2):324–34.

Evangelista, M., Pruyne, D., Amberg, D. C., Boone, C., and Bretscher, A. (2002). Formins direct Arp2/3-independent actin filament assembly to polarize cell growth in yeast. *Nat. Cell Biol.*, 4(1):32–41.

Faulstich, H., Zobeley, S., Rinnerthaler, G., and Small, J. V. (1988). Fluorescent phallotoxins as probes for filamentous actin. *J. Muscle Res. Cell Motil.*, 9(5):370–83.

Fenn, S., Gerhold, C. B., and Hopfner, K.-P. (2011). Nuclear actin-related proteins take shape. *Bioarchitecture*, 1(4):192–5.

Filipenko, N. R. and Waisman, D. M. (2001). The C terminus of annexin II mediates binding to F-actin. *J. Biol. Chem.*, 276:5310–5.

Fischer, H., Robl, I., Sumper, M., and Kröger, N. (1999). Targeting and covalent modification of cell wall and membrane proteins heterologously expressed in the diatom Cylindrotheca fusiformis (Bacillariophyceae). *J. Phycol.*, 35:113–120.

Fujii, T., Iwane, A., Yanagida, T., and Namba, K. (2010). Direct visualization of secondary structures of F-actin by electron cryomicroscopy. *Nature*, 467:724–8.

Galarneau, L., Nourani, a., Boudreault, a. a., Zhang, Y., Héliot, L., Allard, S., Savard, J., Lane, W. S., Stillman, D. J., and Côté, J. (2000). Multiple links between the NuA4 histone acetyltransferase complex and epigenetic control of transcription. *Mol. Cell*, 5(6):927–37.

Galland, R., Leduc, P., Guérin, C., Peyrade, D., Blanchoin, L., and Théry, M. (2013). Fabrication of three-dimensional electrical connections by means of directed actin self-organization. *Nat. Mater.*, 12(5):416–21.

Garcia-Mata, R., Gao, Y.-S., and Sztul, E. (2002). Hassles with taking out the garbage: aggravating aggresomes. *Traffic*, 3(6):388–96.

Gatfield, J., Albrecht, I., Zanolari, B., Steinmetz, M. O., and Pieters, J. (2005). Association of the leukocyte plasma membrane with the actin cytoskeleton through coiled coil-mediated trimeric coronin 1 molecules. *Mol. Biol. Cell*, 16(June):2786–98.

Gibbon, B., Kovar, D., and Staiger, C. (1999). Latrunculin B has different effects on pollen germination and tube growth. *Plant Cell*, 11(12):2349–63.

Glenney, J. R., Geisler, N., Kaulfus, P., and Weber, K. (1981). Demonstration of at least two different actin-binding sites in villin, a calcium-regulated modulator of F-actin organization. *J. Biol. Chem.*, 256(15):8156–61.

Goldschmidt-Clermont, P. J., Machesky, L. M., Doberstein, S. K., and Pollard, T. D. (1991). Mechanism of the interaction of human platelet profilin with actin. *J. Cell Biol.*, 113(5):1081–9.

Gordon, J. L. and Sibley, L. D. (2005). Comparative genome analysis reveals a conserved family of actin-like proteins in apicomplexan parasites. *BMC Genomics*, 6(179):1–13.

Gordon, R., Losic, D., Tiffany, M. A., Nagy, S. S., and Sterrenburg, F. a. S. (2009). The Glass Menagerie: diatoms for novel applications in nanotechnology. *Trends Biotechnol.*, 27(2):116–27.

Gorelik, R., Yang, C., Kameswaran, V., Dominguez, R., and Svitkina, T. (2011). Mechanisms of plasma membrane targeting of formin mDia2 through its amino terminal domains. *Mol. Biol. Cell*, 22(2):189–201.

Grossart, H., Czub, G., and Simon, M. (2006). Algae-bacteria interactions and their effects on aggregation and organic matter flux in the sea. *Env. Microbiol.*, 8(6):1074–84.

Grossart, H. and Simon, M. (2007). Interactions of planktonic algae and bacteria: effects on algal growth and organic matter dynamics. *Aquat. Microb. Ecol.*, 47:163–76.

Gruber, A. and Kroth, P. (2014). Deducing Intracellular Distributions of Metabolic Pathways from Genomic Data. In Sriram, G., editor, *Plant Metab. Methods Protoc. Methods Mol. Biol.*, volume 1083, chapter 12, pages 187–211. Springer, New York, 1083 edition.

Grunt, M., Zárský, V., and Cvrcková, F. (2008). Roots of angiosperm formins: the evolutionary history of plant FH2 domain-containing proteins. *BMC Evol. Biol.*, 8(115):1–19.

Guillard, R. and Ryther, J. (1962). Studies of marine planktonic diatoms. I. Cyclotella nana Hustedt, and Detonula confervacea (cleve) Gran. *Can. J. Microbiol.*, 8(2):229–39.

Gutierrez, R., Lindeboom, J., Paredez, A. R., Emons, A. M. C., and Ehrhardt, D. W. (2009). Arabidopsis cortical microtubules position cellulose synthase delivery to the plasma membrane and interact with cellulose synthase trafficking compartments. *Nat. Cell Biol.*, 11:797 – 806.

Haarer, B. and Lillie, S. (1990). Purification of profilin from Saccharomyces cerevisiae and analysis of profilin-deficient cells. *J. Cell Biol.*, 110:105–14.

Harata, M., Zhang, Y., Stillman, D. J., Matsui, D., Oma, Y., Nishimori, K., and Mochizuki, R. (2002). Correlation between chromatin association and transcriptional regulation for the Act3p/Arp4 nuclear actin-related protein of Saccharomyces cerevisiae. *Nucleic Acids Res.*, 30(8):1743–50.

Harris, E. and Higgs, H. (2004). Actin cytoskeleton: formins lead the way. *Curr. Biol.*, 14:520–2.

Hayes, M. J., Rescher, U., Gerke, V., and Moss, S. E. (2004). Annexin-actin interactions. *Traffic*, 5:571–6.

Heintzelman, M. B. and Enriquez, M. E. (2010). Myosin diversity in the diatom Phaeodactylum tricornutum. *Cytoskeleton*, 67:142–51.

Higaki, T., Sano, T., and Hasezawa, S. (2007). Actin microfilament dynamics and actin side-binding proteins in plants. *Curr. Opin. Plant Biol.*, 10:549–56.

Higgs, H. and Peterson, K. (2005). Phylogenetic analysis of the formin homology 2 domain. *Mol. Biol. Cell*, 16:1–13.

Hliscs, M., Sattler, J. M., Tempel, W., Artz, J. D., Dong, A., Hui, R., Matuschewski, K., and Schüler, H. (2010). Structure and function of a G-actin sequestering protein with a vital role in malaria oocyst development inside the mosquito vector. *J. Biol. Chem.*, 285(15):11572–83.

Holm L, R. P. (2010). Dali server: conservation mapping in 3D. *Nucl. Acids Res.*, 38:545–9.

Holzinger, A. (2010). Jasplakinolide: An Actin-Specific Reagent that Promotes Actin Polymerization. In Gavin, R. H., editor, *Cytoskelet. Methods Protoc.*, volume 586 of *Methods in Molecular Biology*, chapter 4, pages 71–87. Humana Press, Totowa, NJ.

Huang, S., Blanchoin, L., Kovar, D. R., and Staiger, C. J. (2003). Arabidopsis capping protein (AtCP) is a heterodimer that regulates assembly at the barbed ends of actin filaments. *J. Biol. Chem.*, 278(45):44832–42.

Hubberstey, A. V. and Mottillo, E. P. (2002). Cyclase-associated proteins: CAPacity for linking signal transduction and actin polymerization. *FASEB J.*, 16(6):487–99.

Hug, C., Miller, T. M., Torres, M. a., Casella, J. F., and Cooper, J. a. (1992). Identification and characterization of an actin-binding site of CapZ. *J. Cell Biol.*, 116(4):923–31.

Husson, C., Cantrelle, F.-X., Roblin, P., Didry, D., Le, K., Perez, J., Guittet, E., Van Heijenoort, C., Renault, L., and Carlier, M.-F. (2010). Multifunctionality of the βthymosin/WH2 module: G actin sequestration, actin filament growth, nucleation, and severing. *Ann. N.Y. Acad. Sci*, 1194:44–52.

Idei, M., Osada, K., Sato, S., Nakayama, T., Nagumo, T., and Mann, D. G. (2013). Sperm ultrastructure in the diatoms Melosira and Thalassiosira and the significance of the 9 + 0 configuration. *Protoplasma*, 250:833–50.

James, D. E. (1974). *Culturing Algae*. Carolina Biological supply company, Burlington, North Carolina.

Jayo, A. and Parsons, M. (2010). Fascin: a key regulator of cytoskeletal dynamics. *Int. J. Biochem. Cell Biol.*, 42(10):1614–7.

Jiang, Y. and Stillman, D. (1996). Epigenetic effects on yeast transcription caused by mutations in an actin-related protein present in the nucleus. *Genes Dev.*, 10:604–19.

Kabsch, W., Mannherz, H., Suck, D., Pai, E., and Holmes, K. (1990). Atomic structure of the actin: DNase I complex. *Nature*, 347(6):37–44.

Karni, A. and Felder, Y. (2007). Analysis of Biological Networks: Transcriptional Networks - Promoter Sequence Analysis. Technical report.

Kast, D. J. and Dominguez, R. (2011). Arp you ready for actin in the nucleus? *EMBO J.*, 30(11):2097–8.

Khurana, S. and George, S. P. (2008). Regulation of cell structure and function by actin-binding proteins: Villin's perspective. *FEBS Lett.*, 582(14):2128–39.

Kitayama, C. (2003). ForC, a novel type of formin family protein lacking an FH1 domain, is involved in multicellular development in Dictyostelium discoideum. *J. Cell Sci.*, 116(4):711–23.

Kochupurakkal, B. S. and Iglehart, J. D. (2013). Nourseothricin N-acetyl transferase: a positive selection marker for mammalian cells. *PLoS One*, 8(7):e68509.

Komis, G., Apostolakos, P., and Galatis, B. (2002). Hyperosmotic stress induces formation of tubulin macrotubules in root-tip cells of Triticum turgidum: their probable involvement in protoplast volume control. *Plant Cell Physiol.*, 43(8):911–22.

Konopka-Postupolska, D. (2007). Annexins: putative linkers in dynamic membrane-cytoskeleton interactions in plant cells. *Protoplasma*, 230:203–15.

Konopka-Postupolska, D., Clark, G., and Hofmann, A. (2011). Structure, function and membrane interactions of plant annexins: an update. *Plant Sci.*, 181:230–41.

Korenbaum, E., Nordberg, P., Björkegren-Sjögren, C., Schutt, C. E., Lindberg, U., and Karlsson, R. (1998). The role of profilin in actin polymerization and nucleotide exchange. *Biochemistry*, 37(26):9274–83.

Koshino, I., Mohandas, N., and Takakuwa, Y. (2012). Identification of a novel role for dematin in regulating red cell membrane function by modulating spectrin-actin interaction. *J. Biol. Chem.*, 287(42):35244–50.

Kovar, D. R., Drø bak, B. K., and Staiger, C. J. (2000). Maize profilin isoforms are functionally distinct. *Plant Cell*, 12(4):583–98.

Kovar, D. R., Yang, P., Sale, W. S., Drobak, B. K., and Staiger, C. J. (2001). Chlamydomonas reinhardtii produces a profilin with unusual biochemical properties. *J. Cell Sci.*, 114(Pt 23):4293–305.

Kristiansson, E., Thorsen, M., Tamás, M. J., and Nerman, O. (2009). Evolutionary forces act on promoter length: identification of enriched cis-regulatory elements. *Mol. Biol. Evol.*, 26(6):1299–307.

Kröger, N. and Poulsen, N. (2008). Diatoms—From Cell Wall Biogenesis to Nanotechnology. *Annu. Rev. Genet.*, 42:83–107.

Kroth, P. G. (2007). Molecular Biology and the Biotechnological Potential of Diatoms. In León, R., Galván, A., and Fernández, E., editors, *Transgenic Microalgae as Green Cell Factories*, volume 616, chapter 3, pages 23–33. Landes Bioscience and Springer Science+Business Media.

Krügel, H., Fiedler, G., Smith, C., and Baumberg, S. (1993). Sequence and transcriptional analysis of the nourseothricin acetyltransferase-encoding gene nat1 from Streptomyces noursei. *Gene*, 127:127–31.

Kueh, H. Y., Charras, G. T., Mitchison, T. J., and Brieher, W. M. (2008). Actin disassembly by cofilin, coronin, and Aip1 occurs in bursts and is inhibited by barbed-end cappers. *J. Cell Biol.*, 182(2):341–53.

Lancaster, O. and Baum, B. (2014). Shaping up to divide: Coordinating actin and microtubule cytoskeletal remodelling during mitosis. *Semin. Cell Dev. Biol.*, pages 1–7.

Lappalainen, P., Kessels, M. M., Cope, M. J. T., and Drubin, D. G. (1998). The ADF Homology (ADF-H) Domain: A Highly Exploited Actin-binding Module. *Mol. Biol. Cell*, 9:1951–9.

Lauterborn, R. (1896). Untersuchung über Bau, kernteilung und Bewegung der Diatomeen. *Engelmann, Leipzig.*, page 165 pp.

Lázaro-Diéguez, F., Aguado, C., Mato, E., Sánchez-Ruíz, Y., Esteban, I., Alberch, J., Knecht, E., and Egea, G. (2008). Dynamics of an F-actin aggresome generated by the actin-stabilizing toxin jasplakinolide. *J. Cell Sci.*, 121(9):1415–25.

Lee, J. O., Yang, H., Georgescu, M. M., Di Cristofano, a., Maehama, T., Shi, Y., Dixon, J. E., Pandolfi, P., and Pavletich, N. P. (1999). Crystal structure of the PTEN tumor suppressor: implications for its phosphoinositide phosphatase activity and membrane association. *Cell*, 99(3):323–34.

Lemieux, M., Janzen, D., Hwang, R., Roldan, J., Jarchum, I., and Knecht, D. A. (2014). Visualization of the actin cytoskeleton: Different F actin binding probes tell different stories. *Cytoskeleton*, 71(3):157–69.

Li, C.-W. and Volcani, B. (1985). Studies on the Biochemistry and Fine Structure of Silica Shell Formation in Diatoms VIII. Morphogenesis of the Cell Wall in a Centric Diatom, Ditylum brightwellii. *Protoplasma*, 124:10–29.

Li, W. H. (1997). Molecular Evolution. *Sinauer Assoc. Sunderland, MA*.

Lichius, A., Berepiki, A., and Read, N. D. (2011). Form follows function - the versatile fungal cytoskeleton. *Fungal Biol.*, 115(6):518–40.

Liemann, S. and Huber, R. (1997). Three-dimensional structure of annexins. *Cell. Mol. life Sci.*, 53:516–21.

Liemann, S. and Lewit-Bentley, A. (1995). Annexins: a novel family of calcium- and membrane-binding proteins in search of a function. *Structure*, 3:233–7.

Lommer, M., Specht, M., Roy, A.-S., Kraemer, L., Andreson, R., Gutowska, M. a., Wolf, J., Bergner, S. V., Schilhabel, M. B., Klostermeier, U. C., Beiko, R. G., Rosenstiel, P., Hippler, M., and Laroche, J. (2012). Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome Biol.*, 13(R66):1–20.

Lund, P. K., Moats-Staats, B. M., Simmons, J. G., Hoyt, E., D'Ercole, A. J., Martin, F., and Van Wyk, J. J. (1985). Nucleotide sequence analysis of a cDNA encoding human ubiquitin reveals that ubiquitin is synthesized as a precursor. *J. Biol. Chem.*, 260(12):7609–13.

Maciver, S. K. and Hussey, P. J. (2002). The ADF/cofilin family: actin-remodeling proteins. *Genome Biol.*, 3(5):1–12.

Maheswari, U., Mock, T., Armbrust, E. V., and Bowler, C. (2008). Update of the Diatom EST Database: a new tool for digital transcriptomics. *Nucleic Acids Res.*, 37:1–5.

Maheswari, U., Montsant, A., Goll, J., Krishnasamy, S., Rajyashri, K. R., Patell, V. M., and Bowler, C. (2005). The Diatom EST Database. *Nucleic Acids Res.*, 33:D344–7.

Marchler-Bauer, A., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., He, S., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Liebert, C. a., Liu, C., Lu, F., Lu, S., Marchler, G. H., Mullokandov, M., Song, J. S., Tasneem, A., Thanki, N., Yamashita, R. a., Zhang, D., Zhang, N., and Bryant, S. H. (2009). CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res.*, 37(Database issue):D205–10.

Marchler-Bauer, A. and Bryant, S. H. (2004). CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.*, 32(Web Server issue):W327–31.

Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Lu, F., Marchler, G. H., Mullokandov, M., Omelchenko, M. V., Robertson, C. L., Song, J. S., Thanki, N., Yamashita, R. a., Zhang, D., Zhang, N., Zheng, C., and Bryant, S. H. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.*, 39(Database issue):D225–9.

Mathur, J. (2005). Conservation of boundary extension mechanisms between plants and animals. *J. Cell Biol.*, 168:679–682.

Matsuzaki, M., Misumi, O., Shin-i, T., Maruyama, S., Takahara, M., Miyagishima, S.-Y., Mori, T., Nishida, K., Yagisawa, F., Nishida, K., Yoshida, Y., Nishimura, Y., Nakao, S., Kobayashi, T., Momoyama, Y., Higashiyama, T., Minoda, A., Sano, M., Nomoto, H., Oishi, K., Hayashi, H., Ohta, F., Nishizaka, S., Haga, S., Miura, S., Morishita, T., Kabeya, Y., Terasawa, K., Suzuki, Y., Ishii, Y., Asakawa, S., Takano, H., Ohta, N., Kuroiwa, H., Tanaka, K., Shimizu, N., Sugano, S., Sato, N., Nozaki, H., Ogasawara, N., Kohara, Y., and Kuroiwa, T. (2004). Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon merolae 10D. *Nature*, 428(8):653–57.

Mattila, P. and Quintero-Monzon, O. (2004). A high-affinity interaction with ADP-actin monomers underlies the mechanism and in vivo function of Srv2/cyclase-associated protein. *Mol. Biol. Cell*, 15(November):5158–71.

McDowell, J., Huang, S., and McKinney, E. (1996). Structure and evolution of the actin gene family in Arabidopsis thaliana. *Genetics*, 142:587–602.

McFadden, G. and Melkonian, M. (1986). Use of Hepes buffer for microalgal culture media and fixation for electron microscopy. *Phycologia*, 25:551–7.

McGough, A. M., Staiger, C. J., Min, J.-K., and Simonetti, K. D. (2003). The gelsolin family of actin regulatory proteins: modular structures, versatile functions. *FEBS Lett.*, 552(2-3):75–81.

McLaughlin, P., Gooch, J., Mannherz, H., and Weeds, A. (1993). Structure of gelsolin segment 1-actin complex and the mechanism of filament severing. *Nature*, 364:685–92.

Meagher, R., Kandasamy, M., McKinney, E. C., and Roy, E. (2009). Nuclear actin-related proteins in epigenetic control. *Int. Rev. Cell Mol. Biol.*, 6448(09):157–215.

Michelot, A., Berro, J., and Guérin, C. (2007). Actin-filament stochastic dynamics mediated by ADF/cofilin. *Curr. Biol.*, 17(10):825–33.

Michelot, A., Derivery, E., Paterski-Boujemaa, R., Guérin, C., Huang, S., Parcy, F., Staiger, C. J., and Blanchoin, L. (2006). A novel mechanism for the formation of actin-filament bundles by a nonprocessive formin. *Curr. Biol.*, 16(19):1924–30.

Michelot, A., Guérin, C., Huang, S., Ingouff, M., Richard, S., Rodiuc, N., Staiger, C. J., and Blanchoin, L. (2005). The formin homology 1 domain modulates the actin nucleation and bundling activity of Arabidopsis FORMIN1. *Plant Cell*, 17(August):2296–313.

Minoda, A., Saitoh, S., Takahashi, K., and Toda, T. (2005). BAF53/Arp4 homolog Alp5 in fission yeast is required for histone H4 acetylation, kinetochore-spindle attachment, and gene silencing at centromere. *Mol. Biol. Cell*, 16(January):316–27.

Miyagawa Yamaguchi, A., Okami, T., Kira, N., Yamaguchi, H., Ohnishi, K., and Adachi, M. (2011). Stable nuclear transformation of the diatom Chaetoceros sp. *Phycol. Res.*, 59(2):113–9.

Miyahara, M., Aoi, M., Inoue-Kashino, N., Kashino, Y., and Ifuku, K. (2013). Highly efficient transformation of the diatom Phaeodactylum tricornutum by multi-pulse electroporation. *Biosci Biotechnol Biochem.*, 77(4):874–6.

Mock, T. (2013). Adaptation to polar sea ice facilitated by allelic divergence in a psychrophilic diatom. In *Mol. Life Diatoms*, page 32.

Mockrin, S. and Korn, E. (1980). Acanthamoeba profilin interacts with G-actin to increase the rate of exchange of actin-bound adenosine 5'-triphosphate. *Biochemistry*, 19(23):5359–62.

Mohri, K., Vorobiev, S., Fedorov, A. A., Almo, S. C., and Ono, S. (2004). Identification of functional residues on Caenorhabditis elegans actin-interacting protein 1 (UNC-78) for disassembly of actin depolymerizing factor/cofilin-bound actin filaments. *J. Biol. Chem.*, 279(30):31697–707.

Mohseni, M. and Chishti, A. H. (2009). Regulatory models of RhoA suppression by dematin, a cytoskeletal adaptor protein. *Cell Adh. Migr.*, 3(2):191–4.

Morejohn, L., Bureau, T., and Mole-Bajer, J. (1987). Oryzalin, a dinitroaniline herbicide, binds to plant tubulin and inhibits microtubule polymerization in vitro. *Planta*, 172(2):252–64.

Morgan, R. O. and Fernandez, M. P. (2008). Molecular phylogeny and evolution of the coronin gene family. In Clemen, C. S., Eichinger, L., and Rybakin, V., editors, *Coronin Fam. Proteins*, volume 48, chapter 4, pages 41–55. Landes Bioscience and Springer Science+Business Media.

Moriyama, K. and Yahara, I. (2002). Human CAP1 is a key factor in the recycling of cofilin and actin for rapid actin turnover. *J. Cell Sci.*, 115(8):1591–601.

Morrissette, N. and Sibley, L. (2002). Disruption of microtubules uncouples budding and nuclear division in Toxoplasma gondii. *J. Cell Sci.*, 115:1017–25.

Morrissette, N. S., Mitra, A., Sept, D., and Sibley, L. D. (2004). Dinitroanilines bind alpha-tubulin to disrupt microtubules. *Mol. Biol. Cell*, 15(4):1960–8.

Morton, W. M., Ayscough, K. R., and McLaughlin, P. J. (2000). Latrunculin alters the actin-monomer subunit interface to prevent polymerization. *Nat. Cell Biol.*, 2:376–8.

Moss, S. E. and Morgan, R. O. (2004). The annexins. *Genome Biol.*, 5(4):219.

Moustafa, A., Beszteri, B., Maier, U. G., Bowler, C., Valentin, K., and Bhattacharya, D. (2009). Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science*, 324:1724–6.

Muller, J., Oma, Y., Vallar, L., Friederich, E., Poch, O., and Winsor, B. (2005). Sequence and comparative genomic analysis of actin-related proteins. *Mol. Biol. Cell*, 16(December):5736–48.

Mullins, R., Heuser, J., and Pollard, T. (1998). The interaction of Arp2/3 complex with actin: nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *PNAS*, 95(11):6181–6.

Murray, M. and Thompson, W. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.*, 8(19):4321–6.

Muto, M., Fukuda, Y., Nemoto, M., Yoshino, T., Matsunaga, T., and Tanaka, T. (2013). Establishment of a genetic transformation system for the marine pennate diatom Fistulifera sp. strain JPCC DA0580—A high triglyceride producer. *Mar. Biotechnol.*, 15(1):48–55.

Nomura, K. and Ono, S. (2013). ATP-dependent regulation of actin monomer-filament equilibrium by cyclase-associated protein and ADF/cofilin. *Biochem. J.*, 453(2):249–59.

Ochman, H., Gerber, A., and Hartl, D. (1988). Genetic applications of an inverse polymerase chain reaction. *Genetics*, 120(3):621–3.

Oda, T., Iwasa, M., Aihara, T., Maéda, Y., and Narita, A. (2009). The nature of the globular- to fibrous-actin transition. *Nature*, 457(7228):441–5.

Ono, S. (2003). Regulation of actin filament dynamics by actin depolymerizing factor/cofilin and actin-interacting protein 1: new blades for twisted filaments. *Biochemistry*, 42(46):13363–70.

Ono, S. (2013). The role of cyclase-associated protein in regulating actin filament dynamics - more than a monomer-sequestration factor. *J. Cell Sci.*, 126(15):3249–58.

Otomo, T., Tomchick, D. R., Otomo, C., Panchal, S. C., Machius, M., and Rosen, M. K. (2005). Structural basis of actin filament nucleation and processive capping by a formin homology 2 domain. *Nature*, 433(3):488–94.

Özkaynak, E., Finley, D., and Varshavsky, A. (1984). The yeast ubiquitin gene: head-to-tail repeats encoding a polyubiquitin precursor protein. *Nature*, 312:663–6.

Pantaloni, D. and Carlier, M. (1993). How profilin promotes actin filament assembly in the presence of thymosin beta4. *Cell*, 75:1007–14.

Patterson, R. L., van Rossum, D. B., and Gill, D. L. (1999). Store-operated Ca2+ entry: evidence for a secretion-like coupling model. *Cell*, 98(4):487–99.

Paul, A. and Pollard, T. (2008). The role of the FH1 domain and profilin in formin-mediated actin-filament elongation and nucleation. *Curr. Biol.*, 18(1):9–19.

Paul, A. S. and Pollard, T. D. (2009). Review of the mechanism of processive actin filament elongation by formins. *Cell Motil. Cytoskeleton*, 66(8):606–17.

Pettersen, E., Goddard, T., Huang, C., Couch, G., Greenblatt, D., Meng, E., and Ferrin, T. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem*, 25:1605–12.

Pickett-Heaps, J. D. (1991). Post mitotic cellular reorganisation in the diatom Cymatopleura solea: The role of microtubules and the microtubule centre. *Cell Motil. Cytoskeleton*, 292:279–92.

Pickett-Heaps, J. D. and Kowalski, S. (1981). Valve morphogenesis and the microtubule center of the diatom Hantzschia amphioxys. *Eur. J. Cell Biol.*, 25(1):150–70.

Pickett-Heaps, J. D., Schmid, A., and Edgar, L. (1990). The cell biology of diatom valve formation. In Round, F. and Chapman, D., editors, *Prog. Phycol. Res.*, volume 7, pages 1–169. Biopress, Bristol, U.K, 7 edition.

Pickett-Heaps, J. D. and Spurck, T. P. (1982). Studies on kinetochore function in mitosis. I. The effects of colchicine and cytochalasin on mitosis in the diatom Hantzschia amphioxys. *Eur. J. Cell Biol.*, 28(1):77–82.

Poch, O. and Winsor, B. (1997). Who is Who among the Saccharomyces cerevisiae Actin Related Proteins? A Classification and Nomenclature Proposal for a Large Family. *Yeast*, 15:1053–8.

Pollard, T., Blanchoin, L., and Mullins, R. D. (2000). Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Annu. Rev. Biophys. Biomol. Struct.*, 29:545–76.

Pollard, T. and Quirk, S. (1994). Profilins, ancient actin binding proteins with highly divergent primary structures. *Soc. Gen. Physiol. Ser.*, 49:117–28.

Pollard, T. D. and Borisy, G. G. (2003). Cellular motility driven by assembly and disassembly of actin filaments. *Cell*, 112(4):453–65.

Pollock, F. M. and Pickett-Heaps, J. D. (2005). Spatial determinants in morphogenesis: recovery from plasmolysis in the diatom Ditylum. *Cell Motil. Cytoskeleton*, 60(2):71–82.

Pope, B., Way, M., Matsudaira, P., and Weeds, A. (1994). Characterisation of the F-actin binding domains of villin: classification of F-actin binding proteins into two groups according to their binding sites on actin. *FEBS Lett.*, 338:58–62.

Poulsen, N., Chesley, P. M., and Kröger, N. (2006). Molecular genetic manipulation of the diatom Thalassiosira pseudonana (Bacillariophyceae)1. *J. Phycol.*, 42:1059–65.

Poulsen, N. and Kröger, N. (2005). A new molecular tool for transgenic diatoms: Control of mRNA and protein biosynthesis by an inducible promoter-terminator cassette. *FEBS J.*, 272(13):3413–23.

Poulsen, N., Spector, I., Spurck, T. P., Schultz, T. F., and Wetherbee, R. (1999). Diatom Gliding Is the Result of an Actin-Myosin Motility System. *Cell Motil. Cytoskelet.*, 44(1):23–33.

Pruyne, D., Evangelista, M., Yang, C., Bi, E., Zigmond, S., Bretscher, A., and Boone, C. (2002). Role of formins in actin assembly: nucleation and barbed-end association. *Science*, 297(5581):612–5.

Puius, Y. A., Mahoney, N. M., and Almo, S. C. (1998). The modular structure of actin-regulatory proteins. *Curr. Opin. Cell Biol.*, 10(1):23–34.

Rai, S. S. and Wolff, J. (1997). Dissociation of tubulin assembly-inhibiting and aggregation-promoting activities by a vinblastine derivative. *FEBS Lett.*, 416(3):251–3.

Ramalingam, N., Zhao, H., Breitsprecher, D., Lappalainen, P., Faix, J., and Schleicher, M. (2010). Phospholipids regulate localization and activity of mDia1 formin. *Eur. J. Cell Biol.*, 89(10):723–32.

Rana, a. P., Ruff, P., Maalouf, G. J., Speicher, D. W., and Chishti, a. H. (1993). Cloning of human erythroid dematin reveals another member of the villin family. *PNAS*, 90(14):6651–5.

Reimann, B. (1964). Deposition of silica inside a diatom cell. *Exp. Cell Res.*, 34:605–8.

Reimann, B., Lewin, J., and Volcani, B. (1965). Studies on the biochemistry and fine structure of silica shell formation in diatoms. *J. Cell Biol.*, 24:39–55.

Renz, M. and Langowski, J. (2008). Dynamics of the CapG actin-binding protein in the cell nucleus studied by FRAP and FCS. *Chromosom. Res.*, 16(3):427–37.

Riedl, J., Crevenna, A. H., Kessenbrock, K., Yu, J. H., Neukirchen, D., Bista, M., Bradke, F., Jenne, D., Holak, T. A., Werb, Z., Sixt, M., and Wedlich-Soldner, R. (2008). Lifeact: a versatile marker to visualize F-actin. *Nat. Methods*, 5(7):1–8.

Riedl, J., Flynn, K. C., Raducanu, A., Gärtner, F., Beck, G., Bösl, M., Bradke, F., Massberg, S., Aszodi, A., Sixt, M., and Wedlich-Söldner, R. (2010). Lifeact mice for studying F-actin dynamics. *Nat. Methods*, 7(3):168–9.

Rivero, F., Muramoto, T., Meyer, A.-K., Urushihara, H., Uyeda, T. Q. P., and Kitayama, C. (2005). A comparative sequence analysis reveals a common GBD/FH3-FH1-FH2-DAD architecture in formins from Dictyostelium, fungi and metazoa. *BMC Genomics*, 6(28):1–16.

Robinson, R. C., Turbedsky, K., Kaiser, D. a., Marchand, J. B., Higgs, H. N., Choe, S., and Pollard, T. D. (2001). Crystal structure of Arp2/3 complex. *Science*, 294:1679–84.

Rodal, A., Tetreault, J., Lappalainen, P., Drubin, D., and Amberg, D. (1999). Aip1p interacts with cofilin to disassemble actin filaments. *J. Cell Biol.*, 145(6):1251–64.

Romero, S., Le Clainche, C., Didry, D., Egile, C., Pantaloni, D., and Carlier, M.-F. (2004). Formin is a processive motor that requires profilin to accelerate actin assembly and associated ATP hydrolysis. *Cell*, 119(3):419–29.

Rose, A. B. (2002). Requirements for intron-mediated enhancement of gene expression in Arabidopsis. *RNA*, 8(11):1444–53.

Rose, A. B., Elfersi, T., Parra, G., and Korf, I. (2008). Promoter-proximal introns in Arabidopsis thaliana are enriched in dispersed signals that elevate gene expression. *Plant Cell*, 20(3):543–51.

Sagot, I., Klee, S., and Pellman, D. (2002a). Yeast formins regulate cell polarity by controlling the assembly of actin cables. *Nat. Cell Biol.*, 4(January):42–50.

Sagot, I., Rodal, A. a., Moseley, J., Goode, B. L., and Pellman, D. (2002b). An actin nucleation mechanism mediated by Bni1 and profilin. *Nat. Cell Biol.*, 4(8):626–31.

Sambrook, J., Fritsch, E., and Maniatis, T. (2001). *Molecular cloning: a laboratory manual.* Cold Spring Harbour Laboratory, New York.

Sampathkumar, A., Gutierrez, R., McFarlane, H. E., Bringmann, M., Lindeboom, J., Emons, A.-M., Samuels, L., Ketelaar, T., Ehrhardt, D. W., and Persson, S. (2013). Patterning and lifetime of plasma membrane-localized cellulose synthase is dependent on actin organization in Arabidopsis interphase cells. *Plant Physiol.*, 162(2):675–88.

Sato, S., Beakes, G., Idei, M., Nagumo, T., and Mann, D. G. (2011). Novel sex cells and evidence for sex pheromones in diatoms. *PLoS One*, 6(10):e26923.

Schafer, D. and Schroer, T. (1999). Actin-related proteins. *Annu. Rev. Cell Dev. Biol.*, 15:341–63.

Schafer, D. a., Jennings, P. B., and Cooper, J. a. (1996). Dynamics of capping protein and actin assembly in vitro: uncapping barbed ends by polyphosphoinositides. *J. Cell Biol.*, 135(1):169–79.

Schmid, A. (1984). Wall morphogenesis in Thalassiosira eccentrica: Comparison of auxospore formation and the effect of MT-inhibitors. In Mann, D., editor, *Proc. seventh Int. diatom Symp.*, pages 47–70. Koeltz, O., Koenigstein.

Schmid, A. (1987). Morphogenetic forces in diatom cell wall formation. In Bereiter-Hahn, J., Anderson, O., and Reif, W., editors, *Cytomechanics*, pages 183–199. Springer, Berlin Heidelberg New York Tokyo.

Schmid, A. (1994). Aspects of morphogenesis and function of diatom cell walls with implications for taxonomy. *Protoplasma*, 181:43–60.

Schmit, A. (2002). Acentrosomal microtubule nucleation in higher plants. *Int. Rev. Cytol.*, 220:257–89.

Schroer, T. (2004). Dynactin. *Annu. Rev. Cell Dev. Biol.*, 20:759–79.

Schroer, T. and Fyrberg, E. (1994). Actin-related protein nomenclature and classification. *J. Cell Biol.*, 27(6):1777–8.

Schutt, C., Myslik, J., Rozycki, M., Goonesekere, N., and Lindberg, U. (1993). The structure of crystalline profilin-ß-actin. *Nature*, 365:810–6.

Sheterline, P. and Sparrow, J. C. (1994). *Actin. Protein Profile*, volume 1. Oxford University Press, Oxford, 1 edition.

Shimamura, M., Brown, R. C., Lemmon, B. E., Akashi, T., Mizuno, K., Nishihara, N., Tomizawa, K.-I., Yoshimoto, K., Deguchi, H., Hosoya, H., Horio, T., and Mineyuki, Y. (2004). Gamma-tubulin in basal land plants: characterization, localization, and implication in the evolution of acentriolar microtubule organizing centers. *Plant Cell*, 16(1):45–59.

Shu, H. B. and Joshi, H. C. (1995). Gamma-tubulin can both nucleate microtubule assembly and self-assemble into novel tubular structures in mammalian cells. *J. Cell Biol.*, 130(5):1137–47.

Shurety, W., Stewart, N. L., and Stow, J. L. (1998). Fluid-phase markers in the basolateral endocytic pathway accumulate in response to the actin assembly-promoting drug Jasplakinolide. *Mol. Biol. Cell*, 9(4):957–75.

Siebert, P. and Chenchik, A. (1995). An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.*, 23:1087–8.

Siebert, P., Chenchik, A., and Kellogg, D. (1995). The Human GenomeWalker DNA Walking Kit: An new PCR method for walking in uncloned genomic DNA. *Clontechniques*, 10(2):1–3.

Siegel, D. L. and Branton, D. (1985). Partial purification and characterization of an actin-bundling protein, band 4.9, from human erythrocytes. *J. Cell Biol.*, 100(3):775–85.

Sizonenko, G., Karpova, T., Gattermeier, D., and Cooper, J. (1996). Mutational analysis of capping protein function in Saccharomyces cerevisiae. *Mol. Biol. Cell*, 7(January):1–15.

Snapp, E. (2005). Design and use of fluorescent fusion proteins in cell biology. *Curr. Protoc. Cell Biol.*, July:CHAPTER: Unit–21.4.

Spurck, T. P. and Pickett-Heaps, J. D. (1994). The effects of diazepam on mitosis and the microtubule cytoskeleton. *J. Cell Sci.*, 107:2643–51.

Sterck, L., Billiau, K., Abeel, T., Rouzé, P., and de Peer, Y. V. (2012). ORCAE: online resource for community annotation of eukaryotes. *Nat. Methods*, 9.

Szymanski, D. and Cosgrove, D. (2009). Dynamic coordination of cytoskeletal and cell wall systems during plant cell morphogenesis. *Curr. Biol.*, 19(17):R800–R811.

Tamura, K., Peterson, D., Peterson, N., G, S., M, N., and S, K. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.*, 28(10):2731–9.

Tanifuji, G. and Archibald, J. (2010). Actin Gene Family Dynamics in Cryptomonads and Red Algae. *J. Mol. Evol.*, 71(3):169–79.

Tarawneh, K., Anumula, K., and Free, S. (1994). The isolation and characterization of a Neurospora crassa gene (ubi:: crp-6) encoding a ubiquitin-40S ribosomal protein fusion protein. *Gene*, 147:137–40.

Tesson, B., Gaillard, C., and Martin-Jézéquel, V. (2009). Insights into the polymorphism of the diatom Phaeodactylum tricornutum Bohlin. *Bot. Mar.*, 52(2):1–13.

Tesson, B. and Hildebrand, M. (2010a). Dynamics of silica cell wall morphogenesis in the diatom Cyclotella cryptica: substructure formation and the role of microfilaments. *J. Struct. Biol.*, 169(1):62–74.

Tesson, B. and Hildebrand, M. (2010b). Extensive and intimate association of the cytoskeleton with forming silica in diatoms: control over patterning on the meso- and micro-scale. *PLoS One*, 5(12):e14300.

Thomas, C., Tholl, S., Moes, D., Dieterle, M., Papuga, J., Moreau, F., and Steinmetz, A. (2009). Actin bundling in plants. *Cell Motil. Cytoskeleton*, 66(11):940–57.

Tzima, E., Trotter, P. J., Orchard, M. a., and Walker, J. H. (2000). Annexin V relocates to the platelet cytoskeleton upon activation and binds to a specific isoform of actin. *Eur. J. Biochem.*, 267(15):4720–30.

van Attikum, H., Fritsch, O., Hohn, B., and Gasser, S. (2004). Recruitment of the INO80 complex by H2A phosphorylation links ATP-dependent chromatin remodeling with DNA double-strand break repair. *Cell*, 119:777–88.

Van de Meene, A. and Pickett-Heaps, J. (2002). Valve morphogenesis in the centric diatom Proboscia alata sundstrom. *Eur. J. Phycol.*, 38(2):351–63.

Van de Meene, A. M. and Pickett-Heaps, J. D. (2004). Valve morphogenesis in the centric diatom Rhizosolenia setigera (Bacillariophyceae, Centrales) and its taxonomic implications. *Eur. J. Phycol.*, 39(1):93–104.

van Gisbergen, P. a. C., Li, M., Wu, S.-Z., and Bezanilla, M. (2012). Class II formin targeting to the cell cortex by binding PI(3,5)P(2) is essential for polarized growth. *J. Cell Biol.*, 198(2):235–50.

Van Troys, M., Huyck, L., Leyman, S., Dhaese, S., Vandekerkhove, J., and Ampe, C. (2008). Ins and outs of ADF/cofilin activity and regulation. *Eur. J. Cell Biol.*, 87:649–67.

Vidali, L., Rounds, C. M., Hepler, P. K., and Bezanilla, M. (2009). Lifeact-mEGFP reveals a dynamic apical F-actin network in tip growing plant cells. *PLoS One*, 4(5):1–15.

Vinson, V. and Archer, S. (1993). Three-dimensional solution structure of Acanthamoeba profilin-I. *J. Cell Biol.*, 122(6):1277–83.

Voegtli, W., Madrona, A., and Wilson, D. (2003). The structure of Aip1p, a WD repeat protein that regulates cofilin-mediated actin depolymerization. *J. Biol. Chem.*, 278(36):34373–9.

Volkmann, N., Page, C., Li, R., and Hanein, D. (2014). Three-dimensional reconstructions of actin filaments capped by Arp2/3 complex. *Eur. J. Cell Biol.*, 93(5–6):179–83.

Šlajcherová, K., Fišerová, J., Fischer, L., and Schwarzerová, K. (2012). Multiple actin isotypes in plants: diverse genes for diverse roles? *Front. Plant Sci.*, 3(October):226.

Wang, D. and Shaw, G. (1995). of $\beta 1\sigma$II spectrin to brain membranes is mediated by a PH domain, does not require membrane proteins, and coincides with a inositol-1, 4, 5 trisphosphate binding site. *Biochem. Biophys. Res. Commun.*, 217(2):608–15.

Watanabe, N., Madaule, P., and Reid, T. (1997). p140mDia, a mammalian homolog of Drosophila diaphanous, is a target protein for Rho small GTPase and is a ligand for profilin. *EMBO J.*, 16(11):3044–56.

Way, M., Pope, B., Gooch, J., and Weeds, A. (1990). Identification of a region in segment 1 of gelsolin critical for actin binding. *EMBO J.*, 9(12):4103–9.

Werner, D. (1977). *The biology of diatoms*. University of California Press.

Westphal, M., Jungbluth, A., Heidecker, M., Mühlbauer, B., Heizer, C., Schwartz, J.-M., Marriott, G., and Gerisch, G. (1997). Microfilament dynamics during cell movement and chemotaxis monitored using a GFP–actin fusion protein. *Curr. Biol.*, 7:176–83.

Whelan, S. and Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.*, 18:691–9.

Wickstead, B. and Gull, K. (2007). Dyneins across eukaryotes: a comparative genomic analysis. *Traffic*, 8(12):1708–21.

Wieschhaus, A. J., Le Breton, G. C., and Chishti, A. H. (2012). Headpiece domain of dematin regulates calcium mobilization and signaling in platelets. *J. Biol. Chem.*, 287(49):41218–31.

Wordeman, L. and Cande, W. (1990). Cytokinesis by Furrowing in Diatoms. *Ann. New York Acad. Sci.*, pages 252–9.

Wordeman, L., McDonald, K. L., and Cande, W. Z. (1986). The distribution of cytoplasmic microtubules throughout the cell cycle of the centric diatom Stephanopyxis turris: their role in nuclear migration and positioning the mitotic spindle during cytokinesis. *J. Cell Biol.*, 102(5):1688–98.

Xu, J., Casella, J. F., and Pollard, T. D. (1999). Effect of capping protein, CapZ, on the length of actin filaments and mechanical properties of actin filament networks. *Cell Motil. Cytoskeleton*, 42(1):73–81.

Xu, Y., Moseley, J. B., Sagot, I., Poy, F., Pellman, D., Goode, B. L., and Eck, M. J. (2004). Crystal structures of a Formin Homology-2 domain reveal a tethered dimer architecture. *Cell*, 116(5):711–23.

Yamashita, A., Maeda, K., and Maéda, Y. (2003). Crystal structure of CapZ: structural basis for actin filament barbed end capping. *EMBO J.*, 22(7):1529–38.

Yarmola, E. G. and Bubb, M. R. (2006). Profilin: emerging concepts and lingering misconceptions. *Trends Biochem. Sci.*, 31(4):197–205.

Yarmola, E. G., Somasundaram, T., Boring, T. a., Spector, I., and Bubb, M. R. (2000). Actin-latrunculin A structure and function. Differential modulation of actin-binding protein function by latrunculin A. *J. Biol. Chem.*, 275(36):28120–7.

Yin, H. L., Janmey, P. a., and Schleicher, M. (1990). Severin is a gelsolin prototype. *FEBS Lett.*, 264(1):78–80.

Yoshikawa, H., Morimoto, K., Wu, Z., Singh, M., and Hashimoto, T. (2004). Problems in speciation in the genus Blastocystis. *Trends Parasitol.*, 20(6):251–5.

Zaslavskaia, L. and Lippmeier, J. (2000). Transformation of the diatom Phaeodactylum tricornutum (Bacillariophyceae) with a variety of selectable marker and reporter genes. *J. Phycol.*, 36:379–86.

Zhang, C. and Hu, H. (2013). High-efficiency nuclear transformation of the diatom Phaeodactylum tricornutum by electroporation. *Mar. Genomics*, pages 3–6.

Zhang, S.-m., Lv, Z.-y., Zhou, H.-j., Zhang, L.-y., Yang, L.-l., Yu, X., Zheng, H., and Wu, Z.-d. (2008). Characterization of a profilin-like protein from Schistosoma japonicum, a potential new vaccine candidate. *Parasitol. Res.*, 102(6):1367–74.

Zhang, Y., Liu, D., Yang, W., Li, Y., and Yan, H. (2011). A Nonspecific Primer Anchored PCR Technique for Chromosome Walking. *Brazilian Arch. Biol. Technol.*, 54(1):99–106.

Zhang, Y., Vorobiev, S. M., Gibson, B. G., Hao, B., Sidhu, G. S., Mishra, V. S., Yarmola, E. G., Bubb, M. R., Almo, S. C., and Southwick, F. S. (2006). A CapG gain-of-function mutant reveals critical structural and functional determinants for actin filament severing. *EMBO J.*, 25(19):4458–67.

Zigmond, S., Evangelista, M., Boone, C., and Yang, C. (2003). Formin leaky cap allows elongation in the presence of tight capping proteins. *Curr. Biol.*, 13:1820–3.

Abbreviations

aa	Amino acid
BSK-	pBluescript II Sk -
ABP	Actin-binding proteins
ActA	Actin A
ActB	Actin B
ABP	Actin-binding proteins
ADF/cofilin	Actin-depolymerizing factor/cofilin
ALD-DEHYDR-GLU	Aldehyde dehydrogenases glutamic acid active site
ANK-REPEAT	Ankyrin repeat profile
annexin2	Annexin 2 repeat
ARP	Actin-related proteins
ARPAnno	Actin Related Proteins Annotation server
ASB	Actin stabilizing buffer
ATP	Adenosine-5'-triphosphate
bla	Carbenicillin resistance gene
ble	Phleomycin-resistance gene
bp	Base pair
BSA	Bovine serum albumin
C2-TENSIN	Tensin-type C2 domain
CaMV	Cauliflower Mosaic Virus
CAP	Cyclase-associated protein
CapG	Member of the gelsolin family unable
CAPZ	Actin capping protein Z
CARP	C-terminal actin-binding domain (CAP)
CHITINASE-18	Chitinases family 18 active site
ClonNat	Nourseothricin
CTAB	Cetyltrimethyl ammonium bromide
DAPI	4',6-diamidino-2-phenylindole
ddH_2O	Double-distilled water
DDT	Dithiothreitol
DL-DNA	Digested adaptor-ligated genomic DNA fragment
DMSO	Dimethyl sulfoxide
dNTP	Deoxyribonucleotide
DUF2360	Domains of unknown function
EBI	European Bioinformatics Institute
EDTA	Ethylenediamine tetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
EM	Electron microscopy
EST	Expressed sequence tags
EtOH	Ethanol
ExPASy	SIB Bioinformatics Resource Portal
F-actin	Filamentous
FCP	Fucoxanthin chlorophyll binding proteins

FH1	Formin homology 1 domain
FH2	Formin homology 2 domain
FH3	Formin homology 3 domain
Fracy	*Fragilariopsis cylindrus*
G-actin	Globular actin
G418	Geneticin
GBD	GTPase-binding domain
GFP	Green fluorescent protein
H6	C-terminal helix number 6
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HFD	Helically folded domain
HHpred	HMM-HMM comparison
HMMER	Profile hidden Markov models
HMMs	Hidden Markov models
HP	Headpiece
InterPro	Protein Sequence Analysis and classification
IRI	Actin binding motif
JGI	Joint Genome Institute
kbp	Kilo base pair
KIF4	Chromosome-associated kinesin
L-DNA	Adaptor-ligated genomic DNA fragment
LA	Lifeact
MAFFT	Multiple Alignment using Fast Fourier Transform
Mb	Mega bases
MC	Microtubule center
MEGA	Molecular Evolutionary Genetics Analysis software
MQ	'Ultrapure' water of Type 1
mRNA	Messenger ribonucleic acid
MT	Microtubule
MTOC	Microtubule organizing center
MUSCLE	Multiple Sequence Comparison by Log- Expectation
natII	Nourseothricin N-acetyl transferase gene
NCBI	National Center for Biotechnology Information
neo	Neomycin phosphotransferase gene
NLS	Bipartite nuclear localization signal profile
nptII	Neomycin phosphotransferase II gene
P1	Profilin-binding domains
P2	Poly-proline rich region 2
PBS	Phosphate buffered saline
PDB	Protein Data Bank
PEI	Polyethylenimine
PEST region	Proline, glutamic acid, serine, threonine rich region
PH domain	Pleckstrin-homology domain
Phatr	*Phaeodactylum tricornutum*
Pi	Inorganic phosphate
Pipes	Piperazine-N,N'-bis(2-ethanesulfonic acid)
Psemu	*Pseudo-nitzschia multiseries*
PSI-BLAST	Position-Specific Iterated BLAST
PSIPRED	Protein Sequence Analysis Workbench
PTB	Phosphotyrosine-binding domain

PTEN	Phosphatase tensin-type domain
RNA	Ribonucleic acid
RT	Room temperature
SAP	SAF-A/B, Acinus and PIAS
SDV	Silica depositing vesicles
T-particles	Tungsten-particles
T4	T4 ligase
Taq	Omni Polymerase
TEM	Transmission electron microscopy
Thaoc	*Thalassiosira oceanica*
Thaps	*Thalassiosira pseudonana*
TPR	Tetratricopeptide repeat
TRIS	Tris(hydroxymethyl) aminomethane
Ubi1	Ubiquitin 1
Ubi2	Ubiquitin 2
Ubi3	Ubiquitin 3
Ubi4	Ubiquitin 4
UniPortKB	Universal protein database
VH	Villin headpiece
WD40/WD repeats	Beta-transducin repeat
WH2	Wiskott-Aldrich syndrome homology region 2
ZF-FYVE	Zink finger domain named after four proteins: Fab1, YOTB, Vac1, EEA1
ZF-PHD	Plant Homeo Domain

List of Figures

List of Tables

Danksagung

Das größte Dankeschön gilt meinem Mann für seine bedingungslose Unterstützung. Er ist für mich ein Ruhepool und hält mir allzeit den Rücken frei. Meinen drei wundervollen Söhnen Frederik, Leopold und Konstantin danke ich für die für stetige Freude und Abwechslung während meiner Promotion. Meinen Eltern und Schwiegereltern danke ich dafür, dass sie mir immer zur Seite standen und uns unterstützten wenn wir Hilfe benötigten.

Sehr großer Dank gilt meinem Betreuer Prof. D. Menzel für die Möglichkeit, in seiner Arbeitsgruppe meine Promotion anfertigen zu dürfen. Er hat mir neben den zahlreichen Ratschlägen und den weiten methodischen Möglichkeiten die Freiheit zur Entfaltung eigener Forschungsinteressen gegeben und mein eigenständiges Arbeiten sehr unterstützt und gefördert. Dies ermöglichte es mir, bereits während der Promotion viele wertvolle Erfahrungen zu sammeln.

Sehr großer Dank gilt ebenso Prof. P. Kroth und seiner Arbeitsgruppe, im besonderen Ansgar, Dani und Caro. In der Arbeitsgruppe erlernte ich die Transformation von Diatomeen, hatte steten molekular-biologischen und bioinformatischen Austausch und erhielt problemlos Vektoren und Organismen. Für die stete Gastfreundschaft, die herzige Aufnahme, gefolgt von zahlreichen guten Ratschlägen und Aufmunterungen, bis hin zur Übernahme des Zweitgutachtens der Dissertation — Vielen Vielen Dank!

Ganz besonders möchte ich mich bei Herrn Bekir Ülkir für die stundenlangen fruchtbaren Diskussionen bedanken. Bekir Ülkir hat mir die molekularen Methoden näher gebracht und mein Wissen stetig vertieft und erweitert. Durch seine Ideen und Inspirationen hat er zum Gelingen dieser Arbeit beigetragen.

Der ganzen Arbeitsgruppe möchte ich an dieser Stelle sehr für ihre immer präsente und großartige Hilfsbereitschaft danken. Ulla Mettbach besonders für die Erstellung der Ultradünnschnitte von *C. cuspidata* und die anschließende TEM Dokumentation. Claudia Heyme hat mit ihrer Unterstützung bei molekularen Arbeiten zum Gelingen dieser Arbeit beigetragen. Boris Voigt möchte ich dafür danken, dass er sich immer Zeit nahm, stets ein Ohr für mich offen hatte und mir viele hilfreiche Ratschläge wären meiner Promotion gab.

Nicole Poulsen hat es mir ermöglicht in ihren Laboren Experimente durchzuführen und mir sehr bereitwillig Vektoren und Organismen zur Verfügung gestellt. Dies und der stetige wissenschaftliche Austausch mit ihr hat mich sehr voran gebracht und unglaublich motiviert. Ohne ihre Unterstützung wäre diese Arbeit in dem Umfang nicht möglich gewesen.

Bedanken möchte ich mich auch bei den vielen fleißigen Korrekturlesern: Viktoria, Constanze, Fritz, Ansgar, Boris und natürlich Prof. Menzel.

Für die materielle und die ideelle Förderung möchte ich mich bei der Studienstiftung des Deutschen Volkes bedanken. Für die finanzielle Unterstützung zur Erleichterung im Alltag ganz besonders bei der Christiane Nüsslein-Vollhard Stiftung.