Biomolecules Involved in Frustule Biogenesis and Function

Nils Kröger

Abstract

For centuries diatoms have been admired by microscopists for the unusual shapes and intricate, lace-like patterns of their silica-based cell walls. Only comparatively recently, biochemical studies have identified frustule-specific organic biomolecules, of which many are novel and often display remarkably complex chemical structures. This chapter summarizes the current knowledge about the structures and properties of the frustule-associated biomolecules, with a focus on the proteins, and discusses their roles with regards to silica biogenesis and frustule function.

Abbreviations

ACR   Acidic and cysteine-rich
AFIM   Ammonium fluoride insoluble material
AFM   Atomic force microscopy
AFSC   Ammonium fluoride soluble components
AHE   AFIM-HF extract
BHE   Biosilica HF-extract

N. Kröger (✉)
B CUBE, Center for Molecular Bioengineering, TU Dresden, Dresden, Germany

Cluster of Excellence Physics of Life, TU Dresden, Dresden, Germany

Faculty of Chemistry and Food Chemistry, TU Dresden, Dresden, Germany
e-mail: nils.kroeger@tu-dresden.de

Cin   Cingulin
ConA   Concanamycin A
EDTA   Ethylenediaminetetraacetic acid
EE   EDTA extract
GFP   Green fluorescent protein
LCPA   Long-chain polyamines
NMR   Nuclear magnetic resonance
SAP   Silicalemma associated protein
SDS   Sodium dodecyl sulfate
SDV   Silica deposition vesicle
SE   SDS Extract
SEM   Scanning electron microscopy
Sin1   Silicanin-1
TEM   Transmission electron microscopy
VHA   V-type H$^+$-ATPase

1   Frustule-Associated Organic Layers

The diatom cell wall, termed frustule, is an organic–inorganic hybrid material composed of ~90% amorphous, hydrated SiO_2 (silica) and ~10% tightly associated and embedded organic material (Kröger and Poulsen 2008). The inorganic skeleton also contains ions of the metals Al, Fe, Ge, Cu, and Zn (Davis and Hildebrand 2008 and references therein) albeit at rather low concentrations reaching from 0.01% in the case of Zn and Cu to about 1% for aluminum (see Table 1 in Davis and Hildebrand 2008). The role of the inorganic non-silica components on frustule formation and biological function has so far remained largely unexplored. Research on the organic components has been focused mainly on understanding their role in biogenesis of this unique type of cell wall (Hildebrand et al. 2018). During the past 15 years these studies have oncentrated on *Thalassiosira pseudonana* (Fig. 1a),
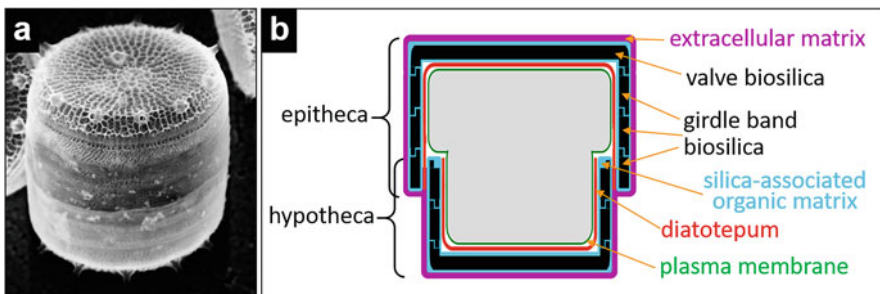

Fig. 1 (**a**) Scanning electron microscopy (SEM) image of the frustule from *T. pseudonana*. (**b**) Schematic diatom cell in cross section showing the biosilica components and organic layers in the frustule. Intracellular compartments are not shown

which grows fairly rapidly to high cell densities, is amenable to genetic modification (Poulsen et al. 2006), and was the first diatom with a completely sequenced genome (Armbrust et al. 2004).

1.1   Arrangement and Biogenesis of the Organic Layers

The general structure of the frustule is shown schematically in Fig. 1b and has been described in detail in Chap. "Structure and Morphogenesis of the Frustule". Electron microscopy imaging revealed the presence of three structurally distinguishable types of organic materials that cover the entire frustule surface both on the proximal and distal side. The organic material on the distal side is termed the extracellular matrix (Fig. 1b, purple). The organic layer between the plasma membrane and the proximal side of the frustule is termed the diatotepum (Fig. 1b, red), which is visible in transmission electron microscopy (TEM) analysis of sectioned cells in regions where the protoplast is detached from the frustule (von Stosch 1981; Crawford 1973). In these regions the diatotepum appears as a well delineated layer that is attached to the frustule only at a few specific sites. Each theca is equipped with its own diatotepum and both layers do not merge at the overlap region, but rather are separated by the girdle bands of the hypovalve (Fig. 1b, red). The third type of organic material in the frustule is the organic casing which completely covers each silica element of the frustule (Fig. 1b, blue) (Volcani 1981). The existence of the organic casing is somewhat controversial, because it is difficult to visualize in TEM sections of intact cells due to its tight association with the silica. However, in TEM images of sections of *C. fusiformis* stained for organic material a distinct layer is visible around each silica element (Reimann et al. 1965; Kröger and Wetherbee 2000). Furthermore, when purified frustules (involving extensive treatment with hydrolytic enzymes) were treated with HF vapor, the silica was completely removed leaving behind a casing that retains the characteristic pattern of the silica (Reimann et al. 1966).

With respect to biogenesis of the organic layers, there is clear evidence that the diatotepum and the extracellular matrix become assembled on the cell surface and are not produced inside the silica deposition vesicles (SDVs), where silica formation takes place (von Stosch 1981; Crawford 1973; van de Poll et al. 1999; Kröger and Wetherbee 2000). To date, biogenesis of the organic casing has remained unclear. Due to its very tight association with the silica it seems likely that it is produced within the SDVs. However, TEM analysis of developing SDVs has not yet provided clear evidence for this scenario. It has been proposed that the organic casing might be derived from the SDV membrane (Pickett-Heaps et al. 1990), yet this seems highly unlikely as the organic casing lacks a lipid bilayer-type morphology in TEM analysis.

1.2   Biochemical Composition of the Organic Layers

Following the discovery of the frustule-associated organic layers, biochemical studies have been undertaken to determine their chemical composition. A large variety of frustule isolation strategies have been developed, which likely have had different effects on the preservation of the organic layers and the presence of contaminating cytoplasmic components. Some of the earliest studies subjected the frustule preparations to exhaustive acid hydrolysis at elevated temperatures, which are conditions that preserve the silica but degrade the main biopolymers (e.g., polysaccharides, proteins, phospholipids). Mixtures of monosaccharides, amino acids, and fatty acids were found in the acid hydrolysates suggesting that the frustule-associated organic material was composed of polysaccharides, proteins, and lipids (for a review see Volcani 1981). Polysaccharides, which are generally highly abundant in algal cell walls, had early on been suspected by microscopists to be the main constituents due to the affinity of the organic layers for polysaccharide-binding dyes. To date, frustule-associated polysaccharides are almost exclusively characterized only regarding their monosaccharide compositions, which is rather complex (Lewin 1955; Ford and Percival 1965a, b; Hoagland et al. 2004; Chiovitti et al. 2005; Willis et al. 2013; Tesson and Hildebrand 2013; Gügi et al. 2015). Whether this reflects the presence of a diverse mixture of polysaccharides, each having a relatively simple monomer composition, or only a few polysaccharides with rather diverse monosaccharide units remains unknown. Recently, a frustule-associated, chitin-based meshwork was identified in *T. pseudonana*, which also contains proteins and is distinct from the chitin fibers that are secreted through specific tube-like openings in the frustules of *Thalassiosira* and *Cyclotella* species (Brunner et al. 2009; Blackwell et al. 1967; Herth and Zungenmaier 1977; Herth 1979; Herth and Barthlott 1979). Using chitin-specific dyes and tagging with green fluorescent protein (GFP), a ring of chitin and a chitin synthase have been localized in the girdle band region of *T. pseudonana* (Durkin et al. 2009; Wustmann et al. 2020). Many other diatom species encode chitin synthase genes (Durkin et al. 2009; Amato and Ferrante 2018), and it seems therefore likely that chitin is a widespread frustule component.

Since proteins and lipids are the main constituents of all cellular components, it could initially not be ruled out that their presence in frustules resulted from contaminating cytoplasmic material (e.g. plasma membrane, plastid and vacuole fragments). Investigations on frustule-associated lipids are very rare, and so far no unusual lipids were found in frustule preparations (Kates and Volcani 1968; Tesson et al. 2008, 2009; Suroy et al. 2014). The few available reports agree that fatty acid bearing lipids are bona fide components of the frustule, yet lipid bilayer-based structures are absent. Therefore, it has remained unresolved how lipids are linked into the organic layers. In contrast to lipids, frustule-associated proteins have been quite extensively studied. This led to the discovery of entirely new protein families, provided the first insight into protein–silica interactions, and stimulated novel hypotheses on silica morphogenesis. The work on frustule-associated lipids and

polysaccharides is still in its infancy, and therefore this review will mainly focus on the role of proteins in frustule formation and function.

The first evidence for the presence of unusual proteins in diatom frustules was obtained by amino acid analysis from Ben Volcani's group more than 50 years ago. They discovered dihydroxyproline and phosphorylated trimethyl-hydroxylysine, which at that time were novel modified amino acids (Nakajima and Volcani 1969, 1970). Furthermore, the standard amino acids serine and threonine were shown to be generally enriched in diatom frustules (Hecky et al. 1973; Swift and Wheeler 1992). About 25 years after Volcani's pioneering work, the first intact proteins from diatom frustules, called frustulins, were isolated and characterized (Kröger et al. 1994). The structures, properties, and possible functions of frustulins will be described below. Before the first diatom genome and transcriptome data became available, the discovery of frustule-associated proteins relied entirely on biochemical analysis of sufficiently pure frustule preparations. A straightforward method developed by the Volcani group, which involves lysis of the cells with glass beads followed by repeated differential centrifugation, yielded remarkably pure frustules from some diatom species (e.g., *Cylindrotheca fusiformis*, *Navicula pelliculosa*) (Nakajima and Volcani 1969, 1970). When *C. fusiformis* frustules were treated with ethylenediamine tetraacetate (EDTA) or sodium dodecylsulfate (SDS), only frustulins were found in the extracts (Kröger et al. 1996). In contrast, the SDS extract of *T. pseudonana* frustules contained many other proteins after the frustules had been extracted with EDTA, urea, and high salt concentrations (Frigeri et al. 2006). It has not yet been systematically investigated by TEM which of the frustule-associated organic layers become fully or partially extracted upon the various extraction procedures. There is indirect evidence from immunofluorescence microscopy that frustulins are part of the extracellular matrix, because they are easily accessible by antibodies. In contrast another family of frustule-associated proteins, termed pleuralins (see below), are only accessible for antibodies after EDTA extraction (Kröger et al. 1997).

Following extraction with EDTA and sodium dodecylsulfate (SDS) at high temperature, the frustules still contain organic components. It is assumed that the vast majority of these are genuine frustule components rather than contaminants, because SDS efficiently solubilizes lipids and proteins. The organic components in SDS-extracted frustules may be highly covalently cross-linked to one another constituting an insoluble organic layer, or they may be partially or fully entrapped inside the silica, or both. It has been speculated that covalent bonds exist between the silica surface, which contains Si-OH and Si-O$^-$ groups, and the functional groups of the organic components. Such covalent linkages could be Si-O-X or Si-X groups with X representing a functional group of the organic component (e.g., CHR_2, COR, NHR, SR). However, in nuclear magnetic resonance (NMR) specroscopy analyses of isolated frustules from *T. pseudonana* and *Stephanopyxis turris* no evidence for covalent bonds between the organic components and the silica was obtained (Christiansen et al. 2006; Jantschke et al. 2015). NMR spectroscopy on intact *N. pelliculosa* cells provided evidence for covalent bonds between silica and the organic components (Kinrade et al. 2002), but no such evidence was found in

T. pseudonana cells in comparable yet even more sensitive NMR experiments (Gröger et al. 2008). Therefore, to date it is generally agreed that the interactions between biomolecules and silica involve exclusively noncovalent bonds.

To extract the organic components from SDS-treated frustules, three different methods were used: HF, NH_4F, and NaOH. Adding liquid, anhydrous HF to dried frustules at 0 °C (boiling point of HF is 19.5 °C) converts silica quantitatively to SiF_4, which escapes as a gas (boiling point of SiF_4: −86 °C). The reaction generates H_2O and thus does not proceed under strictly anhydrous condition, yet HF is in vast excess. During the reaction, O-glycosidically linked sugars are converted to the constituent monosaccharides, phosphoester bonds are cleaved, but N-glycosidic bonds and peptide bonds remain unaffected (Mort and Lamport 1977). Aqueous solutions of HF (30–48%) have also been used effectively for demineralization, but they are prone to hydrolyzing peptide bonds due to the high concentration of water. Nevertheless, peptide data have been obtained from such samples (Frigeri et al. 2006; Nemoto et al. 2020). Aqueous solutions of NaOH completely dissolve silica yielding sodium silicate and at the same time completely hydrolyze proteins. O-glycosidic bonds are stable in NaOH, which has therefore been used to extract polysaccharides from frustules (for example see Chiovitti et al. 2005). The chemically most gentle method for completely dissolving the frustule silica is using a highly concentrated solution of NH_4F at pH 5 (adjusted with HF or HCl), which converts SiO_2 into water soluble SiF_6^{2-}. Peptide bonds, O- and N-glycosidic bonds, and phosphoester bonds are stable under these conditions, enabling the isolation of proteins and probably all other biomolecules in their native chemical structures (Kröger et al. 2002; Poulsen et al. 2003; Poulsen and Kröger 2004). It is assumed that NH_4F treatment extracts biomolecules that are tightly adhered to the silica surface or are embedded within the silica and thus are components of the silica-associated organic matrix (Fig. 1b, blue layer). Biomolecules that are covalently cross-linked into insoluble networks will not get extracted, because the NH_4F solution is unable to cleave such cross-links (e.g., glycosdidic bonds, ester-bonds, isopeptide bonds). Therefore, it is unlikely that material from the extracellular matrix or diatotepum is present in the NH_4F-extract of EDTA- and SDS-extracted frustules as they should have been extracted by these prior treatments.

After dissolving the silica with NH_4F, an insoluble residue is obtained, which is termed AFIM (for *ammonium fluoride insoluble material*). The components in the AFIM originate potentially from any of the three frustule-associated organic layers. They could be meshworks of covalently cross-linked biomolecules, or extremely stable noncovalent assemblies that were not extracted by treatment with EDTA and SDS due to embedment inside the silica. Characterizing the biopolymers that constitute the AFIM is rather difficult due to the insolubility of the material. Solid state ^{13}C-NMR analysis identified β-chitin in *T. pseudonana* in the AFIM (Brunner et al. 2009). After exhaustive acid hydrolysis, complex mixtures of amino acids and a monosaccharides were identified in *T. pseudonana* and other diatoms suggesting that (glyco-) proteins and polysaccharides other than chitin are also present in the AFIM (Tesson and Hildebrand 2013; Kotzsch et al. 2016; Pawolski et al. 2018). Treating the *T. pseudonana* AFIM with anhydrous HF followed by proteomics
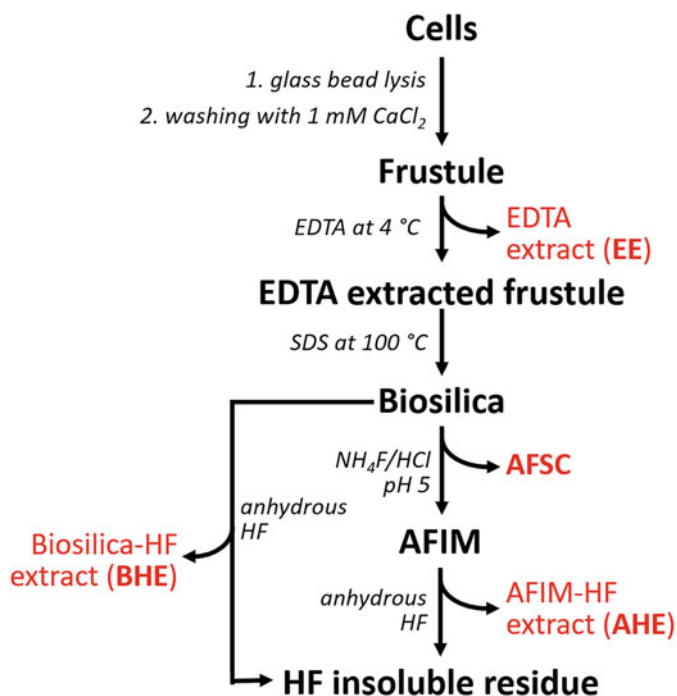
Fig. 2 Extraction of intact proteins and other biomolecules from diatom frustules. The extraction procedures are indicated next to the arrows. AFSC = ammonium fluoride soluble components, AFIM = ammonium fluoride insoluble material

analysis identified 12 proteins (Kotzsch et al. 2016). It was assumed that extraction of these proteins from the AFIM was due to HF-mediated cleavage of cross-links that were based on O-glycosidic and/or phosphodiester bonds (Kotzsch et al. 2016).

The isolation of most of the frustule-associated biopolymers known to date was achieved using the sequential extraction method shown in Fig. 2. In addition to this biochemical approach, frustule-associated proteins have been identified through transcriptomics analysis of synchronized cells, and via bioinformatics approaches (Mock et al. 2008; Scheffel et al. 2011; Shrestha et al. 2012; Brembu et al. 2017). It needs to be stressed that sequence similarity to and/or mRNA co-expression with a known frustule protein should not be regarded as sufficient proof for the frustule association of a protein, for two reasons. Firstly, there is no clear threshold above which similarity in the sequence between two proteins becomes significant for their functional similarity. Secondly, cell cycle synchronization relies on a silicon starvation-replenishment protocol, which not only influences silica biogenesis but also affects the expression of thousands of genes of other metabolic pathways (Brembu et al. 2017). Likewise, not all proteins that are present in extracts obtained from purified frustules are necessarily frustule-associated proteins in vivo, because frustule preparations (as is the case for any subcellular fraction) always contain

contaminating proteins (Frigeri et al. 2006; Kotzsch et al. 2016). As a consequence of the above considerations, at least two independent methods should be applied for unequivocal identification of a frustule-associated protein. For example, proteins identified by proteomics analysis of frustules extracts should only be regarded as bona fide frustule proteins, if they have been located in the frustule through GFP tagging, or immunolocalization, or show an mRNA expression profile that is consistent with frustule biosynthesis, or have high sequence similarity to previously identified frustule proteins. Table 1 lists all currently known frustule-associated proteins that fulfill this criterion. Additionally, it shows from which diatom species long-chain polyamines (LCPA) have been isolated and characterized.

2   Components Involved in Silica Formation

This section will provide an overview of the structures and properties of components that have been shown through in vivo or in vitro experiments to be capable of influencing silica formation. The components include unique proteins, long-chain polyamines (LCPA), and nanopatterned insoluble organic matrices.

2.1   Chemical Structure of LCPA, Silaffins, and Silacidins

LCPA as abundant components in frustules were an entirely unexpected discovery, because they are distinct from proteins, polysaccharides, and lipids that were initially believed to represent the main groups of frustule-associated organic components (see above). LCPA have been isolated from the HF-extracts or NH_4F extracts of EDTA/SDS-cleaned frustule preparations in all diatoms that were analyzed. They are linear oligo-propyleneimine chains that are attached either to diaminopopane or diaminobutane (Kröger and Poulsen 2008; Sumper and Brunner 2008) (Fig. 3a). In the case of diaminobutane as the basis molevcules, LCPA resemble spermine and spermidine, which are ubiquitous polyamines that carry one and two propyleneimine units, respectively. However, diatom LCPA carry between 5–20 propylamine units, show heterogeneous methylation of the N-atoms, and quaternary N-atoms can also be present (Kröger and Poulsen 2008, Sumper and Brunner 2008) (Fig. 3a). Each diatom investigated so far contains a species-specific spectrum of LCPA. Diatom genomes encode unusually high numbers of putative spermine and spermidine synthases. Candidate LCPA synthases have been postulated based on the presence of a signal peptide and unusual bi- and tri-functional polypeptides that have aminopropyltransferase, S-adenosylmethionine decarboxylase, and methyltransferase activities, yet they lack experimental verification (Michael 2011; Nemoto et al. 2020). DNP-ssNMR analysis has demonstrated that LCPA are embedded inside the silica of the diatom *Stephanopyxis turris* rather than being exposed on the surface of the frustule (Jantschke et al. 2015). This result implies that LCPA are present inside SDVs during silica formation, as this is the only conceivable mechanism by which biomolecules could become silica-embedded.

Table 1 Frustule-associated proteins and long-chain polyamines (LCPA) identified in diatoms. The acronyms of the "Frustule extracts" correspond to the nomenclature defined in Fig. 2. The list includes only proteins that have been identified as frustule components through two independent methods. Sp., unidentified species; Ac, *Amphora coffeaeformis*, Cas, *Coscinodiscus asteromphalus*, Ca, *Craspedostauros australis*, Cc, *Cyclotella cryptica*, Cco, *Coscinodiscus concinnus*, Cd, *Chaetoceros didymium*, Cf, *Cylindrotheca fusiformis*, Cg, *Coscinodiscus granii*, Cw, *Coscinodiscus wailesii*, Fs, *Fistulifera solaris*, Na, *Nitzschia angularis*; Np, *Nitzschia palea*, Tp, *Thalassiosira pseudonana*, PPRP, plant pathogenesis-related protein, TPRP, Tetratricopeptide repeat motif containing protein, SPI, serine protease inhibitor, TLSP, Trypsin-like serine protease, STPK, serine/threonine protein kinase (note: this protein is identical to tpSTK1, which is located in the ER rather than the frustule (Sheppard et al. 2010)), HTTRP: protein containing an HSP70 domain and tetratricorepeat domain. References: [1] Kröger et al. 1994, [2] Kröger et al. 1996, [3] Kröger et al. 2002, [4] Poulsen et al. 2003, [5] Kröger et al. 1999, [6] Poulsen and Kröger 2004, [7] Sumper and Brunner 2008, [8] Wenzl et al. 2008, [9] Scheffel et al. 2011, [10] Kotzsch et al. 2016, [11] Kröger et al. 1997, [12] Kröger et al. 2001, [13] Kotzsch et al. 2017, [14] Kröger et al. 2000, [15] Davis and Palenik 2008, [16] Davis et al. 2005, [17] Frigeri et al. 2006, [18] Tesson et al. 2017, [19] Nemoto et al. 2014, [20] Nemoto et al. 2020, [21] Heintze et al. 2020, [22] Buhmann et al. 2014, [23] Lind et al. 1997, [24] Sumper et al. 2005, [25] Pawolski et al. 2018, [26] Sumper and Lehmann 2006

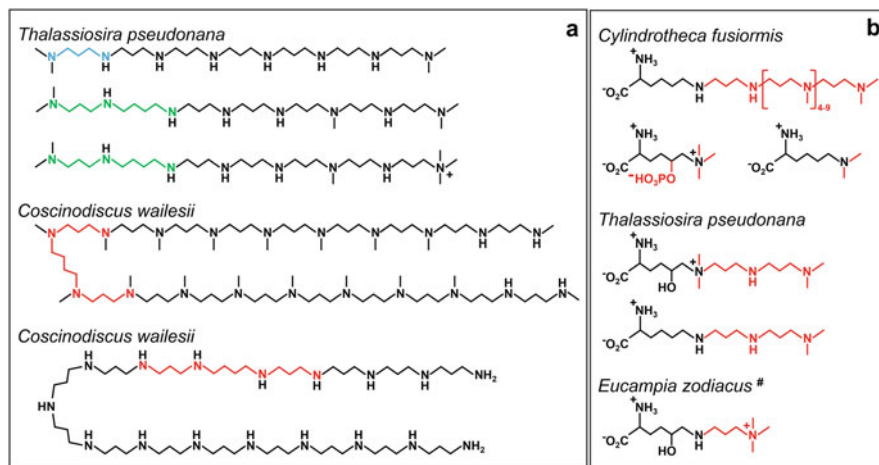| Frustule extract	EE	AFSC	AFIM	BHE	AHE	Other
Cf	Frustulins (α1-3,γ,δ) [1, 2]	Silaffins (natSil1A1,-1A2,-2) [3, 4]		Pleuralins (HEP200,-B,-C) [11] Silaffins (1A,1B,2) [4, 5, 12]		
Tp		Silaffins (tpSil1/2L,-1/2H,-4) [6, 7] Silacidins (A-C) [8] LCPA [6, 24]	Cingulins (Y1, Y4,W1-W3) [9] Silicanin1 [13]	LCPA [14]	Cingulins (Y1-Y4, W3) [10] Silicanin-1 (SiMat7) [10]	p100 [15], p150 [16, 17], SAP1-3 [18] Dymeclin [17], α-Kinase [17], PPRP [17], SPI [17], STPK [17], TPRP [17], TLSP [17], HTTRP [17]
Fs						G7408 [19]
Np						NpSMP1&2 [20]
Cc		LCPA [25]				ccSilicanin1 [21]
Ac			AC3362 [22]			
Ca						FACs [23]
Cd				LCPA [14]		
Nd	ε-Frustulin [2]			LCPA [14]		
Cas, Cw, Cg, Cco		LCPA [26]				

Fig. 3 Examples of oligopropyleneimine bearing molecules from diatoms. (**a**) LCPA; diaminopropane (blue), spermidine (green), and spermine (red) moieties are highlighted in color. (**b**) Polyamine-modified lysines; chemical groups attached to the lysine residues are highlighted in red. Note that most but not all primary, secondary, and tertiary N-atoms in LCPA molecules and the modified lysines will be protonated (Bernecker et al. 2010); (# from Wenzl et al. 2004)

Oligo-propyleneimine chains (in the following abbreviated PA) are also present in silaffins, which is a group of diatom-specific proteins that were found in the HF- and NH_4F-extracts of SDS-extracted frustules (Kröger and Poulsen 2008) (Table 1). Many of the lysine residues in silaffins carry PA modifications at their ε-amino groups. In the case of *C. fusiformis* silaffins, the PA residues carry 6–11 propylamine units (Kröger et al. 2001; Poulsen et al. 2003) while they comprise only two propylamine units in *T. pseudonana* silaffins (Sumper et al. 2007) (Fig. 3b). Generally, a large fraction of a silaffin's lysine residues carry PA modifications, some lysines remain unmodified, and the remainder are dimethylated or trimethylated at the ε-amino group (Kröger et al. 1999, 2001; Poulsen et al. 2003; Poulsen and Kröger 2004; Sumper et al. 2007). Trimethylated lysine residues in silaffins are hydroxylated at the δ-C-atom, and the hydroxyl-group is phosphorylated (Fig. 3b) (Kröger et al. 2001, 2002; Poulsen et al. 2003). Additional phosphorylation sites are serine, threonine, and hydroxyproline residues (Kröger et al. 2002; Poulsen et al. 2003). Further posttranslational modifications include dihydroxylation of proline residues and complex O-glycosylation including sulfated saccharides (Poulsen et al. 2003; Poulsen and Kröger 2004). Besides the extensive addition of functional groups, silaffins are also subject to proteolytic processing in addition to removal of the N-terminal signal peptide for import into the endoplasmic reticulum (Kröger et al. 1999; Poulsen and Kröger 2004). Remarkably, the six silaffins known to date share no significant sequence similarities to each other and are predicted to be intrinsically disordered proteins due to the large predominance of hydrophilic amino acid residues (Scheffel et al. 2011). Therefore, the defining features of a

silaffin are independent of the amino acid sequence and instead comprise (i) the presence of PA-modified lysines, (ii) O-phosphoryl-groups, and (iii) the absence of canonical secondary structure elements. Due to the current scarcity of silaffin sequences and mapped amino acid modification sites, it has remained unknown which features of a silaffin define sites for O-phosphorylation and PA modification. To answer this question, many more silaffin sequences and their tedious biochemical analysis would be required as previously reported for *C. fusiformis* silaffin-1 and *T. pseudonana* tpSil3 (Kröger et al. 2001, 2002; Sumper et al. 2007).

In the NH_4F-extract of EDTA/SDS-cleaned frustules from *T. pseudonana* a second type of phosphoproteins, silacidins, was identified. Silacidins onsist of ~25 amino acid residues of which 35% are serine, 45% are glutamate or aspartate, and no positively charged amino acid residues are present (Wenzl et al. 2008). About 60% of the serine residues are phosphorylated thereby contributing substantially to their extremely high negative charge (Wenzl et al. 2008). Sequence homologs of silacidins have been identified in some but not all diatoms that have been sequenced to date (Kirkham et al. 2017). The *C. fusiformis* silaffin-1 precursor polypeptide contains an N-terminal domain that shows no sequence similarity to silacidins, but is similar regarding the drastic enrichment of serine, aspartate, and glutamate residues. This acidic domain is cleaved off during maturation of the silaffin-generating domain (Kröger et al. 1999). It is currently unknown whether the acidic N-terminal domain becomes phosphorylated and incorporated into the frustule.

Together LCPA, silaffins and siliacidins represent a complete set of charged polyelectrolytes: LCPA are exclusively polycationic, silacidins are exclusively polyanionic, and silaffins are zwitterionic. The net charge of silaffins is dominated by the ratio of cationic propylamine units vs. anionic phosphoryl groups, although sulfated glycan moieties can also have a significant influence (Poulsen et al. 2003). As described in the following section, the charge balance within silaffins as well as in mixtures of the three components is crucial for their ability to accelerate silica formation from silicic acid solutions in vitro.

2.2   Silica Formation by LCPA, Silaffins, and Silacidins

Silica formation assays have been used to investigate in vitro which of the frustule-associated components are capable of influencing the speed by with silica forms in a supersaturated solution of silicic acid. The assays are usually performed at pH 5.5 (buffered by a sodium acetate/acetic acid mixture) as this is presumed to be close to the pH inside the lumen of the SDVs (Vrieling et al. 1999; Shimizu et al. 2001; Yee et al. 2020). At this pH the spontaneous formation of silica by auto-polycondensation of silicic acid is very slow and results in the formation of small silica nanoparticles (diameters <10 nm) after several hours. However, when silaffin natSil1A was added, large silica nanospheres (diameters 400–700 nm) were formed within only 10 min, which demonstrated that this molecule was able to highly accelerate silica formation (Kröger et al. 2002). However, natSil1A cannot be regarded a catalyst for silica polycondensation, because it becomes incorporated

into the silica nanoparticles and thus is consumed in the process. Therefore, natSil1A should be referred to as inducer of silica formation. The silica inducing activity of natSil1A critically depends on the presence of both the polyamine moieties and the phosphoryl groups (Kröger et al. 1999, 2002). It appears that the polypeptide backbone of natSil1A has little if any relevance for silica inducing activity, because rapid formation of large silica nanospheres was also achieved using only LCPA and phosphate in the assay (Kröger et al. 2000; Wenzl et al. 2008). In the absence of inorganic phosphate, LCPA by themselves have very little or no silica formation activity (Wenzl et al. 2008; Kotzsch et al. 2017), whereas mixtures of LCPA and silacidin have high silica formation activity (Wenzl et al. 2008). The efficiency of the silacidin in this system is remarkable, because with respect to phosphate concentration ~1000-fold less silacidin than inorganic phosphate was required to achieve the same silica formation activity (Wenzl et al. 2008). Altogether, the in vitro experiments clearly demonstrated that mixtures of biomolecules with oligo-propyleneimine moieties and polyanionic domains constitute an efficient system for silica formation under biologically relevant conditions (acidic pH, supersaturated silicic acid solution (Martin-Jézéquel et al. 2000; Kumar et al. 2020), time scale of minutes). It should be noted that of all native silaffins known to date, only natSil1A was shown to exhibit intrinsic silica formation activity, whereas the four other silaffins that were tested (natSil2, tpSil1/2H, tpSil1/2L, tpSil3) were inactive on their own. In all these cases the lack of silica forming activity is likely due to a surplus of negative charges within the molecule, which could be clearly demonstrated for natSil2. The glycan chains of natSil2 contain about as many sulfate residues as there are phosphate residues attached to the polypeptide backbone (Poulsen et al. 2003). After complete removal of the glycan chains, the resulting natSil2 derivative, which still contained all phosphate residues, had a silica forming activity that was comparable to the activity of natSil1A (Poulsen et al. 2003). The silica formation activity of deglycosylated natSil2 was eliminated when also the phosphate residues were removed, but was restored by the addition of inorganic phosphate (Poulsen et al. 2003). This set of experiments indicated that the relative ratio of electrical charges between the polycationic oligo-propylamine moieties and the numerous anionic phosphate and sulfate groups determines silica formation activity (Poulsen et al. 2003). This was further confirmed through experiments with mixtures of LCPA and strongly negatively charged silaffins (tpSil3, tpSil1/2H). At a constant concentration of LCPA, low or intermediate concentrations of silaffins enhanced the silica formation activity of the mixture (Poulsen and Kröger 2004) (Fig. 4a). However, from intermediate to high silaffin concentrations, the silica forming activity declined with increasing silaffin concentration (Fig. 4a). The bell-shaped silica vs. LCPA-silaffin diagram can be explained by the presence of LCPA-silaffin clusters that are held together by ionic interactions (Fig. 4a). At low to intermediate silaffin concentration, oligo-propylamine chains are partially accessible to interact with silicic acid molecules (Fig. 4a), which would become concentrated within these clusters thus promoting condensation reactions between the silicic acid molecules. Furthermore, oligo-propyleneimine chains might accelerate silicic acid polycondensation by performing acid-base catalysis (Kröger and Sandhage 2010).
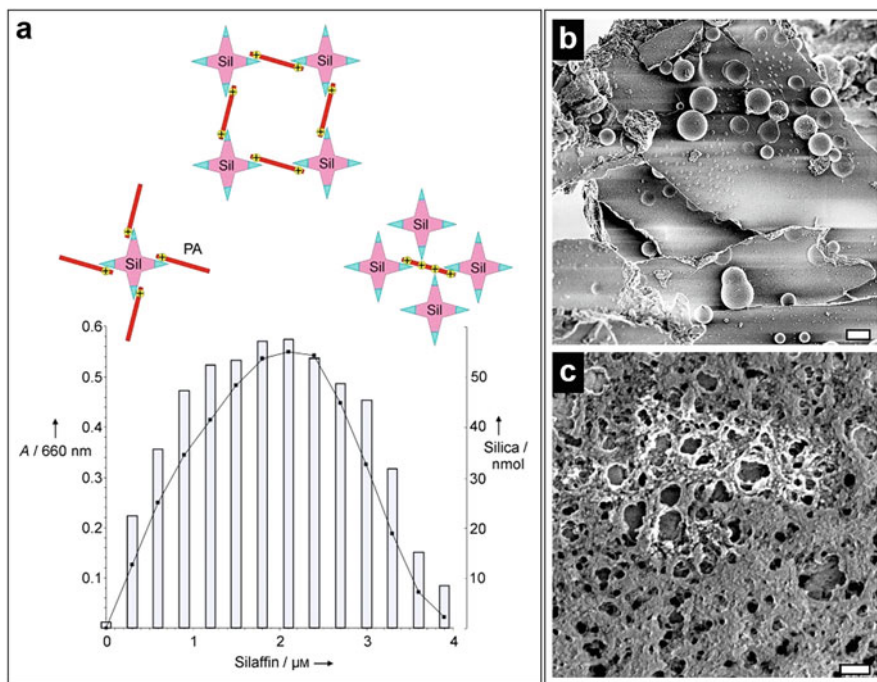
Fig. 4 Aggregation behavior and silica-precipitating activity of LCPA-silaffin mixtures. (**a**) Increasing amounts of tpSil3 were added to a constant concentration of *T. pseudonana* LCPA at pH 5.8. The turbidity of the solution, measured at 660 nm (black dots and black line), increased due to the formation of LCPA-silaffin aggregates. After 10 min incubation with silicic acid the precipitated silica was determined (grey bars). The scheme above the graph shows interpretation of the aggregation behavior of LCPA-tpSil3 mixtures. The positively charged polyamines are shown as red bars and the blue patches denote negatively charged clusters in tpSil3. (**b**, **c**) SEM images of silica formed by mixtures of LCPA with tpSil3 (B) and tpSil1/2H (**c**). Scale bars: 2 µm (**b**), 200 nm (**c**). (**a**) Is from Sumper and Brunner (2008); (**b**, **c**) are from Poulsen and Kröger (2004). Reprinted with permission

At high silaffin concentrations essentially all oligo-propylamine moieties are screened by the negatively charged domains of silaffins (Fig. 4a) prohibiting silicic acid binding and catalysis of polycondensation. Remarkably, the silaffin-to-LCPA stoichiometric ratio does not only influence the quantity, but also the morphology of the silica that is formed. Mixtures of tpSil3 and LCPA produced ~500 nm-thick plates of densely packed <20 nm-sized silica particles and large, polydisperse silica spheres ranging from 900 nm to 4.2 µm in diameter (Fig. 4b) (Poulsen and Kröger 2004). The proportion of silica spheres in the precipitate decreased with increasing tpSil3 concentration (Poulsen and Kröger 2004). Interestingly, a very different silica structure was produced by switching the silaffin, yielding porous silica meshworks in tpSil1/2H-LCPA mixtures (Fig. 4c) (Poulsen and Kröger 2004). The pores were in a biological relevant size range with diameters between 20 and 200 nm (Fig. 4c)

(Poulsen and Kröger 2004). The mechanism by which silaffin-LCPA mixtures guide the formation of silica structures in vitro has so far remained elusive.

It has been hypothesized that electrostatically cross-linked networks of LCPA, silaffins, and silacidins (or other polyanionic molecules) are general components of the silica morphogenesis process inside SDVs (Kröger and Poulsen 2008; Sumper and Brunner 2008). The morphogenesis mechanism may involve spontaneous phase separation (Sumper 2002), which has been shown to occur upon mixing of the components in appropriate stoichiometric ratios (Poulsen et al. 2003). Each diatom species appears to be equipped with a specific set of LCPA and silaffins (Kröger et al. 2000), which may thus allow for the formation of species-specific silica patterns. In vitro experiments using mixtures of LCPA, silaffins and silacidins have so far failed to produce silica structures that match the regularly patterned morphologies of diatom biosilica, although in some cases porous silica was obtained (Poulsen et al. 2003; Poulsen and Kröger 2004). This rather limited success is unsurprising as silica morphogenesis in vivo occurs within an extremely flat micro-container, the SDV, which is delimited by a lipid bilayer. It is reasonable to assume that the confined space of the SDV as well as the membrane composition and properties exert a strong influence on silica morphogenesis.

To understand the roles of LCPA, silaffins and silacidins in diatom silica morphogenesis, the in vitro analyses need to be complemented by in vivo studies. Transcriptomics analysis with a synchronized *T. pseudonana* culture demonstrated that expression of tpSil3 is upregulated during valve formation, whereas tpSil1 expression was downregulated during valve formation and instead coincided with girdle band expression (Shrestha et al. 2012). This result is consistent with the location of the tpSil3-GFP fusion protein, which was found predominantly in the valve and the valve adjacent girdle band when expressed under control of its own promoter (Scheffel et al. 2011). However, it is contradictory to the location of tpSil1-GFP (expressed using the tpSil1 promoter), which was exclusively present in valves (Poulsen et al. 2013). These conflicting data can be reconciled by assuming that tpSil1 is stored in an intermediate compartment before being incorporated into the valve SDV. Alternatively, tpSil1 might never enter the SDV but rather become associated with the valve on the cell surface. LCPA are embedded inside the silica (see above) and thus the highest amounts of LCPA are required during biogenesis of valves, which are the largest frustule building blocks. Consistent with this requirement, the expression of several genes encoding putative enzymes catalyzing oligopropyleneimine biosynthesis (SET domain protein methyltransferases) was enhanced during frustule biosynthesis in *Nitzschia palea* (Nemoto et al. 2020). So far, the expression of the silacidin gene in synchronized *T. pseudonana* cells has not been reported. However, it has been demonstrated that on the protein level the ratio of silacidins vs. tpSil1/2L increases when *T. pseudonana* cells are starved of silicic acid (Richthammer et al. 2011). Based on this result and the observation that already small amounts of silacidins strongly enhance LCPA-mediated silica formation in vitro (see above), it has been suggested that silacidins are critical for frustule formation under conditions of low silicic acid availability (Richthammer et al. 2011). Knocking down the expression pattern of the silacidin gene in *T. pseudonana* using

antisense RNA had significant effects on the size and silica content of the cells. The frustules of the knockdown mutants contained 21% less silica per surface area than the cells of the "control strain" (i.e. transformant expressing the antibiotic resistance gene but not the silacidin antisense RNA), which is consistent with a role of silacidin in silica biogenesis (Kirkham et al. 2017). Additionally, the knockdown mutants exhibited a 25% increase in valve diameter and a 2.2-fold higher cell volume, and thus were both wider and more elongated than wild type cells (Kirkham et al. 2017). It has been speculated that silacidins promote nutrient uptake by being involved in maintaining a favorable surface-to-volume ratio of diatoms cells (Kirkham et al. 2017).

2.3   Proteins of the SDV Membrane

Only omparatively recently, the first SDV membrane proteins have been discovered through two complementary approaches. In one approach chitin-free insoluble organic matrices from *T. pseudonana* were subjected to proteomics analysis and one of the identified proteins was Silicanin-1 (Sin1, formerly SiMat7) (Kotzsch et al. 2016). GFP-tagging in combination with time lapse confocal fluorescence microscopy demonstrated that Sin1 is co-located with valve and girdle band SDVs (Kotzsch et al. 2017). In the complementary approach, transcriptomics analysis of a synchronized *T. pseudonana* culture revealed numerous genes with silaffin-like expression patterns, which were then subjected to a bioinformatics screen to identify proteins with both an N-terminal ER-import signal peptide and a single transmembrane domain (Tesson et al. 2017). One of the resulting proteins, termed silicalemma-associated protein 1 (SAP1), was GFP-tagged and found to be associated with valve and girdle band SDVs through confocal fluorescence microscopy (Tesson et al. 2017). In the following the current knowledge on the SDV membrane proteins is summarized.

Sin1 is a predicted type-1 transmembrane protein with an N-terminal signal peptide for co-translational ER import, a ~40 kDa luminal domain, a single transmembrane helix, and a short C-terminal domain (20 amino acids) (Kotzsch et al. 2017). Sin1 is highly conserved among all diatoms from which transcriptome or genome data are available (sequence identity 46–66%). *T. pseudonana* encodes a protein with 67% sequence identity to Sin1, which was termed Sin2 (Kotzsch et al. 2017) and is also associated with valve and girdle band SDVs (Kotzsch and Kröger, unpublished data). The C-termini of Sin1 and Sin2 are devoid of any sequence motifs with known function, and thus it is unclear whether they interact with components in the cytoplasm. In the SDV lumen Sin1 will likely be engaged in forming supramolecular assemblies, because its luminal domain is prone to homo-multimerization at mildly acidic conditions and to interact with LCPA in vitro (Kotzsch et al. 2017). During SDV development, Sin1 becomes proteolytically processed and the luminal domain incorporated into the silica-associated organic matrix (Kotzsch et al. 2017). The fate of the cleaved-off membrane anchor and C-terminal domain is unclear (Kotzsch et al. 2017). Sin1 knockout mutants of

T. pseudonana exhibited subtle yet quantifiable changes in valve but not girdle band morphology, and the silica content of the frustules was substantially reduced (67–77% of wild type frustules) (Görlich et al. 2019). The reduction in frustule silica content in Sin1 knockout mutants is consistent with the observation that the luminal domain of Sin1 enhances the silica formation activity of LCPA. The mechanism by which Sin1 influences valve morphology remains to be determined.

SAP1 is also a type-1 transmembrane protein, but it does not share sequence similarity with silicanins (Tesson et al. 2017). The luminal part of SAP1 is extremely rich in serine (15.5%), and the cytosolic C-terminus (28 amino acids) does not contain any sequence motifs of known function (Tesson et al. 2017). Two other type-1 transmembrane proteins with similar sequences, termed SAP2 and SAP3, are encoded in the *T. pseudonana* genome and additional homologs were identified in 11 other diatom species (Tesson et al. 2017). GFP-tagging revealed that SAP3 is also located in valve and girdle band SDVs whereas SAP2-GFP, surprisingly, resided in the cytoplasm (Tesson et al. 2017). During SDV maturation SAP1 was proteolytically processed and the luminal part incorporated into the mature frustule. It is not yet known whether the luminal domains of SAP1 and SAP3 become part of the insoluble organic matrices like Sin1, or become associated with the silica in a different way. It has been speculated that the serine residues of SAPs are phosphorylated and might influence silica morphogenesis through interaction with the oligo-propyleneimine chains of LCPA and silaffins in the SDV lumen (Tesson et al. 2017). Although there is so far no experimental evidence for such interactions and for SAPs being phosphoproteins, knockdown mutants of SAP1 and SAP3 did exhibited alterations in silica morphology (Tesson et al. 2017).

SDVs are acidic compartments (see above), and recent evidence indicates that a vacuolar type H^+-ATPase (in the following termed VHA) is responsible for this. GFP-tagging revealed that VHA is located not only in the vacuole of *T. pseudonana* but also in valve and girdle band SDVs (Yee et al. 2020). When a small molecule inhibitor of VHA, concanamycin A (ConA), was applied in vivo, multiple SAP1-GFP bearing, spherical vesicles accumulated in the cytoplasm, and VHA-GFP was present throughout the cytoplasm with a slight accumulation in the mid-cell region of dividing cells where the valve SDVs are located in unperturbed cells (Yee et al. 2020). Only some of the SAP1-GFP bearing vesicles of ConA-treated cells accumulated the acidotropic dye PDMPO indicating that their acidification was severely compromised (Yee et al. 2020). It appears that SDVs are unable to form in ConA-treated cells, yet electron microscopy data of sectioned cells are lacking to corroborate this assumption. At low ConA concentrations, some cells were still able to divide, but the morphology of the valve was severely compromised (Yee et al. 2020). Altogether, these data show that the H^+-pumping function of VHA is essential for vesicle-mediated delivery of proteins to the SDV, for acidification of the SDV, and possibly also for SDV assembly. Furthermore, VHA being also located in the vacuolar membrane raises the question, whether vesicle-mediated transport exists between the SDVs and the vacuole. This question challenges the prevailing view that SDV biogenesis is exclusively fueled by Golgi-derived vesicles.
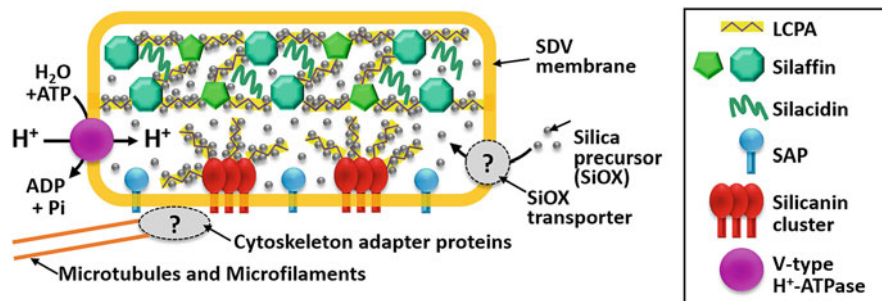
Fig. 5 Model of an SDV. The model assumes the formation of a silaffin-LCPA-silacidin phase self-assembled through polyvalent ionic interactions. Silica precursor molecules may accumulate within this phase through interactions with the oligo-propyleneimine moieties. Within the silaffin-LCPA-silacidin phase the accumulated precursor molecules are rapidly converted into patterned silica structures (not shown). The membrane-bound silicanin clusters bind LCPA molecules thereby either constituting an independent silica forming phase or interacting with the silaffin-LCPA-silacidin phase. The H^+-pumping V-type ATPase is involved in establishing an acidic pH milieu in the SDV. As interaction partners of SAPs are currently unknown, it can only be speculated that they might influence silica morphogenesis via binding to silica-forming phases that are composed of any of the other SDV components

Assuming that the molecules described in Sect. 2 are located in the SDV during silica biogenesis, and taking into account the experimental evidence on their possible functions, a picture of the molecular organization in the SDV has emerged (Fig. 5). This picture includes also information presented in other chapters on association of the cytoskeleton (microfilaments, microtubules) with the SDV (see Chap. "Structure and Morphogenesis of the Frustule") and the presence of a yet unknown silica precursor (SiOX) in the cytoplasm (see Chap. "Silicic Acid Uptake and Storage by Diatom"). Association of microtubules and microfilaments with the SDV requires yet unknown cytoskeleton adapter proteins that bind to both the cytoskeleton filament and the cytoplasmic face of the SDV membrane (Fig. 5). None of the three classes of SDV membrane proteins known to date (silcanins, SAPs, VHA) bears known microfilament or microtubule binding domains. Therefore, these proteins may either not interact with the cytoskeleton, or bind to cytoskeleton adapter proteins. Alternatively, they may directly bind microfilaments and microtubules through novel interaction domains. Regarding transport of the silica precursor SiOX into the SDV lumen, it is currently unknown what molecule is transported and by what mechanism (passive diffusion, facilitated diffusion, or active transport; see Chap. "Silicic acid Uptake and Storage by Diatom"). It seems certain that SiOX is a silicic acid derivative different from monosilicic acid $Si(OH)_4$, which is taken up by the cell and stabilized in the cytoplasm by an unknown mechanism (Hildebrand 2008). It is anticipated that in addition to the presumed cytoskeleton adapter proteins and the SiOX transporter, the SDV will contain yet unknown components in the lumen, the SDV membrane, and on the cytoplasmic

surface. The most straightforward way to identify such components would require developing a method for SDV isolation, which has so far not been achieved.

3   Insoluble Organic Matrices

Most, if not all, biominerals leave behind an organic residue when the mineral phase is gently, yet completely dissolved. Diatoms silica is no exception. After dissolving SDS/EDTA extracted *T. pseudonana* frustules with a mildly acidic solution of ammonium fluoride, three structurally different types of organic material have been identified: meshworks, microrings, and microplates (Fig. 6a–c). In the following, the compositions and properties of these ammonium fluoride insoluble materials (AFIM) and their possible role in silica biogenesis are described.

3.1   Biochemical Compositions of the AFIM

The meshworks in the *T. pseudonana* AFIM are mainly composed of chitin and contain a small proportion of associated proteins (Brunner et al. 2009). The microrings and microplates are mainly composed of proteins (~75%) but also contain a substantial amount of carbohydrates (~25%) with a complex monosaccharide composition (Kotzsch et al. 2016). Organic matrices were structurally characterized also in *N. pelliculosa* (Reimann et al. 1966), *C. wailesii* (Scheffel et al. 2011), *Cyclotella cryptica* (Pawolski et al. 2018), *Coscinodiscus radiatus*, *Stephanopyxis turris*, *Nitzschia curvilineata*, *Triceratium dubium*, *Amphora salina* (all in Tesson and Hildebrand 2013). Characteristic for these matrices is a fairly regular nano- and microarchitecture that is akin to the biosilica patterns (Fig. 6a–e). For example, in *T. pseudonana* the stripe patterns of the microrings are identical to the nonporous margins of the girdle bands (Fig. 6b), and the dot patterns in microplates correspond to the fultoportula patterns in the valves (Fig. 6c).

Biochemical analysis of the AFIM has been severely hampered by their insolubility, even in strongly denaturing aqueous solvents. Therefore, their compositions have so far only been characterized using either complete acid hydrolysis or incomplete digestion by anhydrous HF. Amino acid composition analysis has been performed for *T. pseudonana* (Brunner et al. 2009; Kotzsch et al. 2016) and *C. cryptica* (Pawolski et al. 2018), and analysis of monosaccharide composition for *T. pseudonana* (Kotzsch et al. 2016), *C. radiatus*, *N. curvilineata*, *A. salina*, and *T. dubium* (all in Tesson and Hildebrand 2013). The monosaccharide compositions were rather complex, yet mannose was always a dominant component. The amino acid compositions of the insoluble organic matrices from *T. pseudonana* and *C. cryptica* shared a marked predominance of glycine and serine, which is consistent with an early report on the amino acid composition of diatom frustules (Hecky et al. 1973).

A preparation containing microrings and microplates from *T. pseudonana* is, so far, the only insoluble organic matrix material from which peptide sequences have
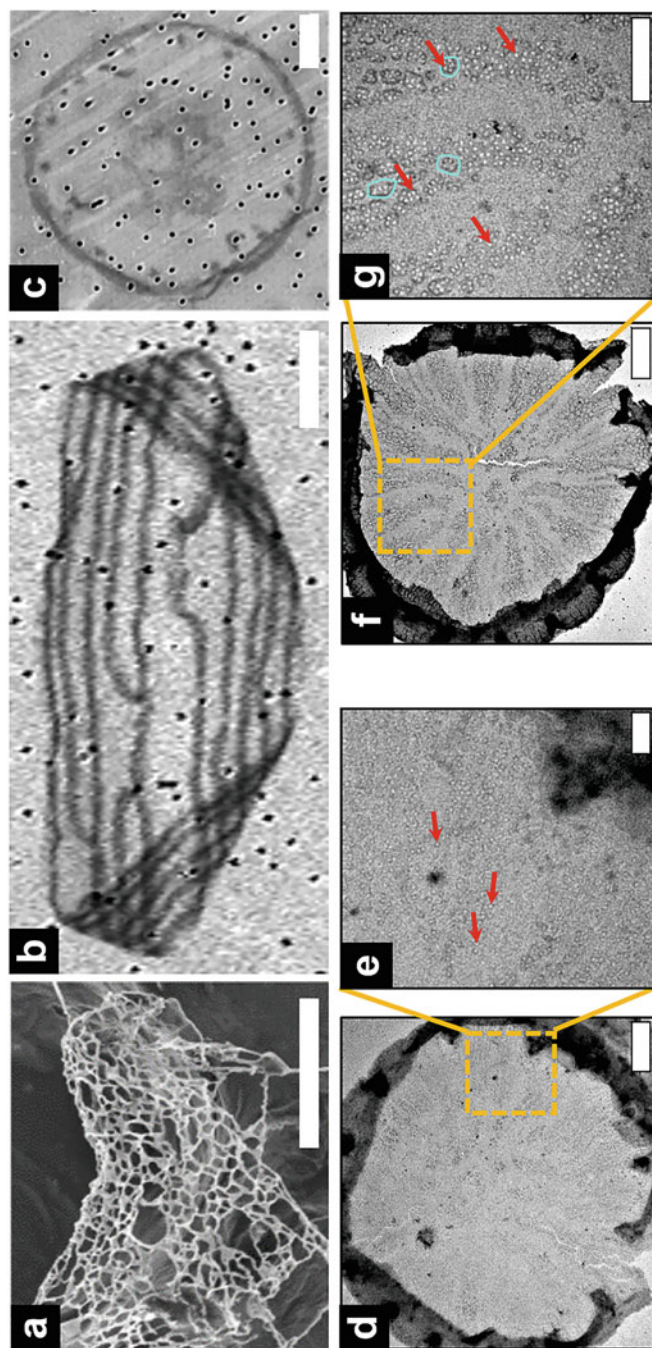
Fig. 6 Ammonium fluoride insoluble organic matrices. (**a**) Chitin meshwork, (**b**) microring, and (**c**) microplate from *C. cryptica*, and (**e**) a higher magnification of the region of the microplate that is framed in (**d**). (**f**) A *C. cryptica* microplate after incubation with silicic acid in the presence of the soluble organic components. (**g**) Detail of the region framed in (**f**). The red arrows point to some electron translucent circles resembling cribrum pores; the turquoise rings highlight some areola like pores. (**a–c**) SEM images, (**d–g**) TEM images. Scale bars: 1 μm (**a–d**), 250 nm (**e**, **g**), 500 nm (**f**). (Modified from Brunner et al. 2009 (**a**), Scheffel et al. 2011 (**b**), Kotzsch et al. 2016 (**c**), Pawolski et al. 2018 (**d–g**)). Reprinted with permission

been be obtained. This lead to the identification of 12 proteins, of which seven had not been previously identified, and were denoted SiMat1-7 (for silica matrix; Kotzsch et al. 2016). Of the five known proteins, four were cingulins (CinY1, CinY2, CinY3, CinW3), which were previously identified in a bioinformatics search for silaffin-like proteins (Scheffel et al. 2011). The bioinformatics search retrieved two additional cingulins (CinW1, CinW2), for which no peptide data were obtained (Scheffel et al. 2011). One protein, silaffin tpSil1, had previously been identified in the NH$_4$F-extracts (Poulsen and Kröger 2004). GFP-tagging revealed that all cingulins are specifically associated with girdle bands in vivo and are present in the microrings, thus confirming that miocrorings are girdle band-specific organic matrices (Scheffel et al. 2011). The amino acid sequence of SiMat1 exhibited features that are very similar to Y-type cingulins, and GFP-tagging demonstrating that this protein is located in girdle bands/microrings (Kotzsch et al. 2016). Therefore, SiMat1 has been renamed CinY4 (Kotzsch et al. 2016). SiMat7 was localized to the SDV membrane through GFP-tagging. It was renamed silicanin-1 (Sin1), and its structures and possible functions have been described above.

3.2 Role of Insoluble Organic Matrices in Silica Formation

The structural resemblance of the organic microrings and microplates with the biosilica of girdle bands and valves, respectively, has spurred the hypothesis that insoluble organic matrices might act as templates in silica morphogenesis (Scheffel et al. 2011). Therefore, in vitro mineralization experiments were performed during which surface-adsorbed insoluble organic matrices were exposed to metastable solutions of silicic acid at pH 5.5 in the presence or the absence of soluble components. The filamentous pattern in the microrings from *T. pseudonana* was enhanced after mineralization due to preferential silica deposition along the filaments (Scheffel et al. 2011). However, no porous silica patterns were formed. Adding synthetic polyamines accelerated silica formation, but had no influence on the silica pattern (Scheffel et al. 2011). In TEM analysis, microplates isolated from *C. cryptica* revealed patterns reminiscent of the hierarchical patterns of cribrum and areola pores in the valve biosilica (Fig. 6d, e) (Pawolski et al. 2018). In AFM analysis patterns of both pores and globules were observed in the regions where TEM revealed only pore-like patterns (Pawolski et al. 2018). When incubated with silicic acid, silica bearing, hierarchical pore patterns were formed that covered the entire surface in about 35% of the microplates. When the ammonium fluoride soluble material (i.e., proteins and LCPA) was included in the mineralization reaction the amount of microplates containing silicified hierarchical pore patterns over the entire surface doubled (Fig. 6f, g) (Pawolski et al. 2018). This enhancement was exclusively due to the presence of LCPA, whereas the soluble proteins did not contribute to morphogenesis of the pore patterns, but seemed to enhance the amount of deposited silica (Pawolski et al. 2018). This study demonstrated that ammonium fluoride soluble organic components and the insoluble organic matrices act synergistically in mineral morphogenesis in vitro. However, only relatively weakly silicified structures were

obtained that reconstituted the hierarchical pore patterns but no other features of the *C. cryptica* biosilica (costae, fultoportulae).

It is currently unknown whether the chitin meshwork identified in *T. pseudonana* (Brunner et al. 2009) and the chitin-containing margin of *C. cryptica* microplates (Pawolski et al. 2018) promotes silica formation. Studies with secreted chitin fibers from *T. pseudonana*, which are distinct from the chitin meshwork, were unable to accelerate silica formation within biologically relevant times and pH (Spinde et al. 2011). NMR studies revealed that Si-OH groups of silica that formed spontaneously by auto-polycondensation after 24 h form a hydrogen bond specifically with the C6-OH group rather than C3-OH group of the N-acetylglucosamine unit of chitin (Spinde et al. 2011). This result suggests that in vivo mineralization of the chitin meshwork would require the activity of silica-forming components (e.g., LCPA, silaffins), but chitin itself can establish strong binding to the silica via numerous hydrogen bonds. It is unclear, whether chitin meshwork, micropates, or microrings are involved in silica morphogenesis in vivo. Such a role would require the insoluble organic matrix to be present inside the SDV during silica biogenesis, which remains to be investigated.

4   Components Involved in Frustule Function

The function of the frustule is subject to much speculation. It has been argued that biosynthesis of a silica-based cell wall requires much less energy than would be required for a purely organic cell wall (Raven 1983), and would provide increased resistance to parasites (Raven and Waite 2004). The frustule promotes the enzymatic conversion of HCO_3^- to CO_2, by acting as a pH buffer, thus aiding in inorganic carbon acquisition (Milligan and Morel 2002). Due to the porous patterns, frustules are able to influence the propagation of light into the cell thus providing protection against harmful UV light (De Tommasi et al. 2018) and aiding in the acquisition of light for photosynthesis (Goessling et al. 2018). Due to the intricate architectures with ridges and pores, diatom frustules have extraordinary mechanical properties exhibiting the highest specific strength of any known biological material (Aitken et al. 2016). It has been hypothesized that the frustule acts as armor in the defense against zooplankton predators like copepods (Hamm et al. 2003). At the same time, the architecture, rigidity, and chemistry of the frustule also provides challenges. It imposes a unique mode of cell division and growth (see Chap. "Cellular Hallmarks and Regulation of the Diatom Cell Cycle") during which the contact zone of the two halves of the frustule need to be repeatedly loosened and re-sealed. Furthermore, most aqueous environments are highly under-saturated regarding silica formation (2 mM at pH 7), and thus the frustule is thermodynamically driven to dissolve. The following section deals with the biomolecules that appear to play a role in coping with the challenges imposed by the chemistry and architecture of the frustule.

4.1   Preventing Silica Dissolution

The dissolution of silica from frustules is enhanced by bacteria due to the secretion of exoproteases (Bidle and Azam 1999). It seems evident that this effect was caused by the degradation of frustule-associated proteins. One of the likely targets of bacterial exoproteases are the frustulins, which are Ca^{2+}-binding glycoproteins that appear to be present in all diatom species (Kröger et al. 1994, 1996; Nemoto et al. 2020). The characteristic feature of frustulins is the presence of multiple ACR domains (for *a*cidic, *c*ysteine-*r*ich) each containing a conserved C-$^E/_Q$-G-D-C-D motif. In *C. fusiformis*, frustulins form a seemingly continuous coat around all parts of the biosilica (Kröger et al. 1997), and their Ca^{2+}-affinity is required for attachment to the frustule. Frustulins are acidic proteins and thus might impede silica dissolution by neutralizing OH^- ions near the silica surface that would otherwise catalyze the cleavage of Si-O-Si bonds. Although this scenario is speculative, the constitutive synthesis and incorporation of frustulins into the frustule (Kröger et al. 1994) is consistent with a neutralizing function.

Other proteins that encase large parts of frustules are the tyrosine-rich protein AC3362 from *A. coffeaeformis* (Buhmann et al. 2014) and frustule-associated glycoproteins, termed FACs, from *Craspedostauros australis* (formerly *Stauroneis decipiens*) (Lind et al. 1997). The sequences of FACs have not yet been determined, but studies with monoclonal antibodies against FACs pointed to their potential role in adhesion to surfaces (Lind et al. 1997) (see Chap. "Diatom Adhesion and Motility"). Currently there is no evidence for FACs playing a role in protecting the frustule from silica dissolution. AC3362 is a basic rather than acidic protein, and it is part of the ammonium fluoride insoluble matrix in *A. coffeaeformis* (Buhmann et al. 2014). The protein does not appear to be present in any other diatoms (Buhmann et al. 2014), and is thus unlikely to be involved in a function that is as generic as protecting frustules from silica dissolution. In this context, it is interesting to speculate that silaffins and LCPA, which are believed to be general components of the biosilica-forming machinery (see above), might also play a role in impeding silica dissolution. Mixtures of silaffins and LCPA strongly promote silica formation from silicic acid in vitro (see above), which means that they decrease silica solubility. It will be challenging to prove a component's involvement in preventing silica dissolution in vivo as such an activity would be difficult to distinguish from promoting silica biogenesis.

4.2   Regulating Epitheca-Hypotheca Contact in the Overlap Region

A functionally very interesting part of the frustule is the region where the epitheca overlaps the hypotheca (Fig. 7a). During the cell cycle the composition of the overlap region is subjected to continuous stepwise change as explained in the following. Immediately before cell division, each theca is equipped with its maximum number of girdle bands and the terminal girdle bands (called pleural bands)

face each other (Fig. 7a). During cell division (i.e. cytokinesis and valve biogenesis, see Fig. 1 in Chap. "Cellular Hallmarks and Regulation of the Diatom Cell Cycle") the contact between the pleural bands is broken and the pleural band of each theca establishes contact with the edge of the newly formed valve (Fig. 7a). Subsequently, the phase of cell growth starts with adding the first girdle band to the new valve. This requires breaking the pleural band-valve contacts and establishing contact between the new girdle band and the pleural band of the epitheca. Cell growth continues in a stepwise manner by successively adding new girdle bands to the hypotheca, and each time the epitheca pleural band exchanges its contact partner in the growing girdle band region of the hypotheca (Fig. 7b). These modes of cell division and growth ensure that the protoplast is always fully encased by silica. It appears that at any stage of frustule development epitheca and hypotheca are glued together, because the frustule remains intact even when the proptoplasma is removed by extraction with detergent (e.g. 2% SDS at 95 °C). Upon acid treatment, girdle bands become completely separated from each other and from the valve, suggesting that connections between the silica elements in an intact frustule are achieved via biopolymers that are sensitive to acid hydrolysis, like proteins and polysaccharides.
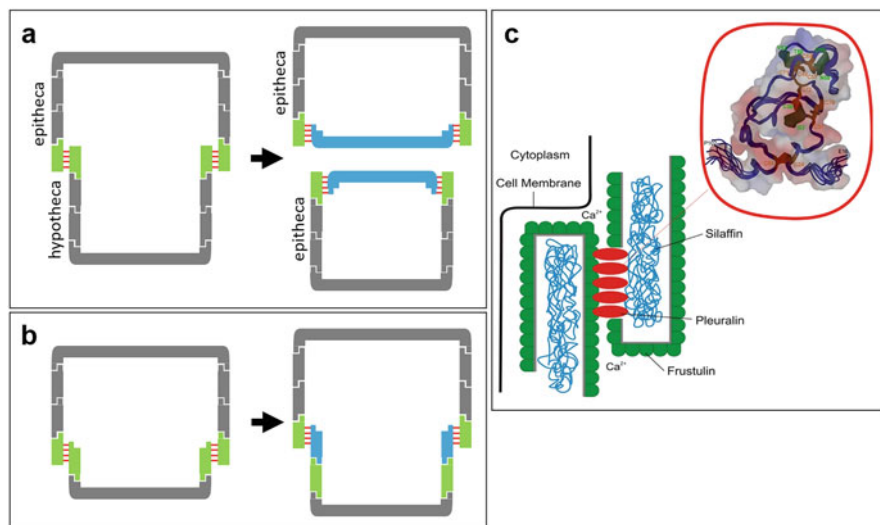

Fig. 7 Contacts between epitheca and hypotheca during the cell cycle. (**a, b**) Schematic diatom cells are shown in cross section. For simplicity the protoplasts are not shown. Frustule components are shown in gray and in color. Green: girdle bands in the overlap region before exocytosis of a new girdle band/valve; blue: newly formed girdle band/valve.; red lines: molecules mediating epitheca–hypotheca contact. (**a**) Valve formation, (**b**) girdle band formation. (**c**) Model of the arrangement and interactions of pleuralins, frustulins, and silaffins in the overlap region of the *C. fusiformis* frustule (from De Sanctis et al. 2016). Reprinted with permission. The 3D structure of the PSCD4 domain from pleuralin-1 is shown on the upper right corner. Disulfide bonds are shown in orange and the electrostatic surface potentials are red for negative charge and blue for positive charge

Indeed, chitin and specific types of frustule-associated proteins, p150 and pleuralins, have been localized to the overlap region in diatoms using fluorescent probes and antibodies, respectively (Durkin et al. 2009; Wustmann et al. 2020; Kröger et al. 1997; Kröger and Wetherbee 2000; Davis et al. 2005). Chitin was particularly abundant in the overlap region when the cells were stationary due to nutrient limitation or silicon starvation (Durkin et al. 2009; Wustmann et al. 2020). The chitin in the overlap region is likely associated with p150, which is an acidic, 150 kDa protein that contains three chitin binding sites (Davis et al. 2005). Like for chitin, the abundance of p150 strongly increases upon silicon starvation, but in contrast to chitin it is not restricted to the overlap region and appears to be associated with all girdle bands (Davis et al. 2005). The increased abundance of chitin and p150 in the overlap region of stressed, stationary cells suggests that these two components are necessary to stabilize the epitheca–hypotheca contact. Stationary cells neither produce girdle bands nor valves and thus no temporary loosening of the epitheca–hypotheca contact is necessary until the conditions become favorable for cell growth.

Pleuralins are pleural band-associated proteins that have so far only been identified in *C. fusiformis* (Kröger et al. 1997). Their characteristic feature is the presence of multiple PSCD domains (each ~100 amino acids), each containing five disulfide bridges with proline, serine, cysteine, and aspartate constituting ~50% of the amino acids (Kröger et al. 1997). Due to the high proline content, the

PSCD domain is almost completely devoid of canonical secondary structures and has a loosely packed structure (Fig. 7c) (De Sanctis et al. 2016). The localization of pleuralins in nutrient stressed or stationary cells has not been investigated, but their locations at different stages of the cell cycle has been determined using immunogold electron microscopy. This revealed that pleuralins are located only at the pleural bands of the epitheca and completely absent from the hypotheca (Kröger and Wetherbee 2000). During cell division, pleuralins are secreted into the extracellular space between the two daughter cells from where they bind to the pleural bands of the hypotheca (Kröger and Wetherbee 2000). At this stage the hypotheca is primed to become the epitheca of one of the daughter cells as soon as a new valve has been produced (Fig. 7a). It was hypothesized that the presence of pleuralins in *C. fusiformis* frustules signals completion of the girdle band region and that the theca is now prepared to function as an epitheca (Kröger and Wetherbee 2000). NMR studies have shown that a recombinant PSCD domain interacts with frustulins and silaffins (De Sanctis et al. 2016), which suggests that pleuralins may be directly involved in connecting the pleural bands of the epitheca with the hypotheca (Fig. 7c). The observations that Ca^{2+} ions induce conformational changes in the PSCD domains (De Sanctis et al. 2016) and frustulin binding to the frustule requires Ca^{2+} (Kröger et al. 1994) spur the speculation that Ca^{2+} concentration changes in the overlap region trigger reversible binding and unbinding between epitheca and hypotheca. However, at this point it is not even known whether pleuralin–frustulin interactions occur across the overlap region or only within the same theca.

4.3   Achieving High Mechanical Performance

A key function of the frustule is mechanical protection against zooplankton grazers, which is believed to provide diatoms with a selective advantage over other algae being - at least partially - responsible for their superior productivity in the natural environment (Hamm et al. 2003). Indeed, under laboratory conditions the copepod *Parvocalanus crassirostris* ingested *T. weissflogii* cells that had frustules with strongly reduced silica content with a 2–3 times higher rate than the control *T. weissflogii* cells that had ~3 times higher silica content (Liu et al. 2016). In the natural environment, diatoms have to compromise between silica content of the cell wall and the need to stay afloat in the photic zone (for planktonic diatoms) as well as restricting the time for silicic acid uptake to still enable high duplication rates. Considering the importance of the armor function of the cell wall, one would expect genes that are crucial for the mechanical performance of frustules to be highly conserved among all diatoms. In agreement with this is the assumed function of frustulins (silica dissolution inhibitors) and their seemingly ubiquitous presence among the diatoms (see above). Sin1, which is present in the SDV membrane and incorporated into the biosilica (see above), is the only other highly conserved frustule protein known so far. *Sin1* knockout mutants of *T. pseudonana* exhibited normal growth rates under laboratory conditions, yet they had reduced biosilica content (70% of wild type) and displayed relatively mild morphological aberrations in their valves (Görlich et al. 2019). Nanoindentation measurements revealed a 3.5-fold reduced strength and stiffness of the frustules of the *sin1* knockout mutants (Görlich et al. 2019). Further investigations are required to clarify if the reduced silica content alone or also the morphological aberrations in the valve are responsible for the decreased mechanical performance of the frustules in *sin1* knockout mutants.

5   Conclusion

Diatom cell walls have unique architectures ranging from the patterns of nano- to microsized pores, ribs, tubes, and spines to the bipartite arrangement of their shells. However, they are not the only organisms capable of making and shaping silica. Silica is the taxonomically most widespread biomineral, being produced by a plethora of organisms from all eukaryotic supergroups (Marron et al. 2016). Biosilica formation also exist among prokaryotes (Motomura et al. 2016). This raises the question, whether all silica forming organisms have inherited this physiological capability from a common ancestor, or whether it has evolved multiple times independently. It is yet too early to answer this question, as the molecular underpinnings of silica biogenesis have so far only been studied in sufficient detail in diatoms and sponges. Remarkably, both types of organisms employ completely different sets of proteins for this process (Otzen 2012), yet they do share the use of LCPA (Matsunaga et al. 2007). Sponges also contain polyamine-modified biomacromolecules, but it is not known whether these are proteins or other acid hydrolysis-sensitive biopolymers (Matsunaga et al. 2007). Furthermore, polyamine-

modified lysines, identical to those from silaffins, have been found in the silica scales of a haptophyte alga (Durak et al. 2016). Linear and branched oligo-propyleneimine chains are also present in bacteria and archaea species, yet none of these are known to produce silica (Knott 2009). It is therefore tempting to speculate that the first silica forming organisms adapted preexisting LCPA molecules, which originally had other functions, for the novel purpose of controlling cellular silica deposition in a silicic acid-saturated environment. Different groups of silica forming organisms might then have independently evolved proteins to produce silica morphologies for adaptation to different life style requirements. To enable thorough investigation of the evolutionary history of silica formation in diatoms, in depth proteomics analyses of diatom frustules and SDVs are required along with the genome sequences from a wide variety of silica forming organisms. Furthermore, the enzyme machineries for LCPA biosynthesis need to be identified and compared among diatoms and other silica forming organisms.

Why is it important to study biological silica formation, and why in diatoms? As is detailed in several other chapters of this book, diatoms are key primary producers in the oceans, and their metabolism and proliferation strictly depends on their ability to biosynthesize silica. Therefore, understanding the mechanism of diatom silica formation and function, and the factors that influence them, are crucial for understanding fluxes of matter and energy in ocean ecosystems. Furthermore, the mechanisms for genetically controlled biomorphogenesis of mineralized structures is, so far, largely unanswered for any biomineral forming organism. Diatoms produce biominerals with complex structures that exhibit species-specific characteristics on the nano, meso-, and microscale (from ~10 nm up to ~100 μm), yet they are biologically simple (single-celled) and readily amenable for genetic manipulation. This makes them prime model systems for studying a cellular machinery that translates a genetic blueprint into a 3D mineral architecture. Finally, the relatively young research field of diatom nanotechnology, develops methods for utilizing diatom silica with its exceptional structures and properties for many applications ranging from energy harvesting to biomedicine (Losic 2017). A deep understanding of the silica formation mechanisms would enable the generation of engineered diatoms that produce silica-based materials with tailor-made structure and properties (Kröger and Brunner 2014; Kröger et al. 2018). This would represent an entirely new paradigm for materials production through a carbon-neutral and renewable biotechnological process.

References

Aitken ZH, Luo S, Reynolds SN, Thaulow C, Greer JR (2016) Microstructure provides insights into evolutionary design and resilience of *Coscinodiscus sp.* frustule. Proc Natl Acad Sci U S A 113: 2017–2022

Amato A, Ferrante MI (2018) Chitin synthases in diatoms. Ann Proteomics Bioinform 1:1001

Armbrust EV et al (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science 306:79–86

Bernecker A, Wieneke R, Riedel R, Seibt M, Geyer A, Steinem C (2010) Tailored synthetic polyamines for controlled biomimetic silica formation. J Am Chem Soc 132:1023–1031

Bidle KD, Azam F (1999) Accelerated dissolution of diatom silica by marine bacterial assemblages. Nature 397:508–512

Blackwell J, Parker KD, Rudall KM (1967) Chitin Fibres of the diatoms *Thalassiora fluviatilis* and *Cyclotella cryptica*. J Mol Biol 28:383–385

Brembu T, Chauton MS, Winge P, Bones AM, Vadstein O (2017) Dynamic responses to silicon in *Thalasiossira pseudonana* - identification, characterisation and classification of signature genes and their corresponding protein motifs. Sci Rep 7:4865

Brunner E, Richthammer P, Ehrlich H, Paasch S, Simon P, Ueberlein S, van Pée K-H (2009) Chitin-based organic networks: an integral part of cell wall biosilica in the diatom *Thalassiosira pseudonana*. Angew Chem Int Ed 48:9724–9727

Buhmann MT, Poulsen N, Klemm J, Kennedy MR, Sherrill CD, Kröger N (2014) A tyrosine-rich cell surface protein in the diatom *Amphora coffeaeformis* identified through transcriptome analysis and genetic transformation. PLoS One 9:e110369

Chiovitti A, Harper RE, Willis A, Bacic A, Mulvaney P, Wetherbee R (2005) Variations in the substituted 3-linked mannans closely associated with the silicified walls of diatoms. J Phycol 41: 1154–1161

Christiansen SC, Hedin N, Epping JD, Janicke MT, del Amo Y, Demarest M, Brzezinski M, Chmelka BF (2006) Sensitivity considerations in polarization transfer and filtering using dipole–dipole couplings: implications for biomineral systems. Solid State Nucl Mag Res 29: 170–182

Crawford RM (1973) The organic component of the cell wall of the marine diatom *Melosira nummuloides* (Dillw.) C. Ag. Br Phycol J 8:257–266

Davis AK, Hildebrand M (2008) A self-propagating system for Ge incorporation into nanostructured silica. Chem Commun 2008:4495–4497

Davis AK, Palenik B (2008) Characterization of a modular, cell-surface protein and identification of a new gene family in the diatom *Thalassiosira pseudonana*. Protist 159:195–207

Davis AK, Hildebrand M, Palenik B (2005) A stress-induced protein associated with the girdle band region of the diatom *Thalassiosira pseudonana* (Bacillariophyta). J Phycol 41:577–589

De Sanctis S, Wenzler M, Kröger N, Malloni WM, Sumper M, Deutzmann R, Zadravec P, Brunner E, Kremer W, Kalbitzer H-R (2016) PSCD domains of pleuralin-1 from the diatom *Cylindrotheca fusiformis*: NMR structures and interactions with other biosilica-associated proteins. Structure 24:1178–1191

De Tommasi E, Congestri R, Dardano P, De Luca AC, Managò S, Rea I, De Stefano M (2018) UV-shielding and wavelength conversion by centric diatom nanopatterned frustules. Sci Rep 8: 16285

Durak GM, Taylor AR, Walker CE, Probert I, de Vargas C, Audic S, Schroeder D, Brownlee C, Wheeler GL (2016) A role for diatom-like silicon transporters in calcifying coccolithophores. Nat Commun 7:10543

Durkin CA, Mock T, Armbrust EV (2009) Chitin in diatoms and its association with the cell wall. Euk Cell 8:1038–1050

Ford CW, Percival E (1965a) The carbohydrates of *Phaeodactylum tricornutum*. Part I. preliminary examination of the organism, and characterisation of low molecular weight material and of a glucan. J Chem Soc 1965:7035–7041

Ford CW, Percival E (1965b) Carbohydrates of *Phaeodactylum tricornutum*. Part II. A sulphated glucuronomannan. J Chem Soc 1965:7042–7046

Frigeri LG, Radabaugh TR, Haynes PA, Hildebrand M (2006) Identification of proteins from a cell wall fraction of the diatom *Thalassiosira pseudonana*: insights into silica structure formation. Mol Cell Proteomics 5:182–193

Goessling JW, Su Y, Cartaxana P, Maibohm C, Rickelt LF, Trampe ECL, Walby SL, Wangpraseurt D, Wu X, Ellegaard M, Kühl M (2018) Structure-based optics of centric diatom frustules: modulation of the *in vivo* light field for efficient diatom photosynthesis. New Phytol 219:122–134

Görlich S, Pawolski D, Zlotnikov I, Kröger N (2019) Control of biosilica morphology and mechanical performance by the conserved diatom gene Silicanin-1. Commun Biol 2:245

Gröger C, Sumper M, Brunner E (2008) Silicon uptake and metabolism of the marine diatom *Thalassiosira pseudonana*: solid-state ^{29}Si NMR and fluorescence microscopic studies. J Struct Biol 161:55–63

Gügi B, Le Costaouec T, Burel C, Lerouge P, Helbert W, Bardor M (2015) Diatom-specific oligosaccharide and polysaccharide structures help to unravel biosynthetic capabilities in diatoms. Mar Drugs 13:5993–6018

Hamm CE, Merkel R, Springer O, Jurkojk P, Maier C, Prechtel K, Smetacek V (2003) Architecture and material properties of diatom shells provide effective mechanical protection. Nature 421:841–843

Hecky RE, Mopper K, Kilham P, Degens ET (1973) The amino acid and sugar composition of diatom cell walls. Mar Biol 19:323–331

Heintze C, Formanek P, Pohl D, Hauptstein J, Rellinghaus B, Kröger N (2020) A new view into the silica deposition vesicles of diatoms. BMC Mater 2:11

Herth W (1979) The site of beta-chitin fibril formation in centric diatoms. II. The chitin-forming cytoplasmic structures. J Ultrastruct Res 68:16–27

Herth W, Barthlott W (1979) The site of beta-chitin fibril formation in centric diatoms. I. Pores and fibril formation. J Ultrastruct Res 68:6–15

Herth W, Zungenmaier P (1977) Ultrastructure of the chitin fibrils of the centric diatom *Cyclotella cryptica*. J Ultrastruct Res 61:230–239

Hildebrand M (2008) Diatoms, biomineralization processes, and genomics. Chem Rev 108:4855–4874

Hildebrand M, Lerch SJL, Shrestha RP (2018) Understanding diatom cell wall silicification – moving forward. Front Mar Sci 5:125

Hoagland KD, Rosowski JR, Gretz MR, Roemer SC (2004) Diatom extracellular polymeric substances: function, fine structure, chemistry, and physiology. J Phycol 29:537–566

Jantschke A, Koers E, Mance D, Weingarth M, Brunner E, Baldus M (2015) Insight into the supramolecular architecture of intact diatom biosilica from DNP-supported solid-state NMR spectroscopy. Angew Chem Int Ed 54:15069–15073

Kates M, Volcani BE (1968) Studies on the biochemistry and fine structure of silica shell formation in diatoms. Lipid components of the cell walls. Z Pflanzenphysiol 60:19–29

Kinrade SD, Gillson A-ME, Knight CTG (2002) Silicon-29 NMR evidence of a transient hexavalent silicon complex in the diatom *Navicula pelliculosa*. J Chem Soc Dalton Trans 2002:307–309

Kirkham AR, Richthammer P, Schmidt K, Wustmann M, Maeda Y, Hedrich R, Brunner E, Tanaka T, van Pée K-H, Falciatore A, Mock T (2017) A role for the cell-wall protein silacidin in cell size of the diatom *Thalassiosira pseudonana*. ISME J 11:2452–2464

Knott JM (2009) Biosynthesis of long-chain polyamines by crenarchaeal polyamine synthases from *Hyperthermus butylicus* and *Pyrobaculum aerophilum*. FEBS Lett 583:3519–3524

Kotzsch A, Pawolski D, Milentyev A, Shevchenko A, Scheffel A, Poulsen N, Shevchenko A, Kröger N (2016) Biochemical composition and assembly of biosilica-associated insoluble organic matrices from the diatom *Thalassiosira pseudonana*. J Biol Chem 291:4982–4997

Kotzsch A, Gröger P, Pawolski D, Bomans PHH, Sommerdijk NAJM, Schlierf M, Kröger N (2017) Silicanin-1 is a conserved diatom membrane protein involved in silica biomineralization. BMC Biol 15:65

Kröger N, Brunner E (2014) Complex-shaped microbial biominerals for nanotechnology. WIREs Nanomed Nanobiotechnol 6:615–627

Kröger N, Poulsen N (2008) Diatoms - from cell wall biogenesis to nanotechnology. Annu Rev Genet 42:83–107

Kröger N, Sandhage KH (2010) From diatom biomolecules to bioinspired syntheses of silica- and titania-based materials. MRS Bull 35:122–126

Kröger N, Wetherbee R (2000) Pleuralins are involved in theca differentiation in the diatom *Cylindrotheca fusiformis*. Protist 151:263–273

Kröger N, Bergsdorf C, Sumper M (1994) A new calcium binding glycoprotein family constitutes a major diatom cell wall component. EMBO J 13:4676–4683

Kröger N, Bergsdorf C, Sumper M (1996) Frustulins: domain conservation in a protein family associated with diatom cell walls. Eur J Biochem 239:259–264

Kröger N, Lehmann G, Rachel R, Sumper M (1997) Characterization of a 200 kDa diatom protein that is specifically associated with a silica-based substructure of the cell wall. Eur J Biochem 250:99–105

Kröger N, Deutzmann R, Sumper M (1999) Polycationic peptides from diatom biosilica that direct silica nanosphere formation. Science 286:1129–1132

Kröger N, Deutzmann R, Bergsdorf C, Sumper M (2000) Species-specific polyamines from diatoms control silica morphology. Proc Natl Acad Sci U S A 97:14133–14138

Kröger N, Deutzmann R, Sumper M (2001) Silica precipitating peptides from diatoms: the chemical structure of silaffin-1A from *Cylindrotheca fusiformis*. J Biol Chem 276:26066–26070

Kröger N, Lorenz S, Brunner E, Sumper M (2002) Self-assembly of highly phosphorylated silaffins and their function in biosilica morphogenesis. Science 298:584–586

Kröger N, Dubey NC, Kumari E (2018) Immobilization of proteins on diatom biosilica. In: Diatom nanotechnology. Royal Society of Chemistry, London, pp 126–149

Kumar S, Rechav K, Kaplan-Ashiri I, Gal A (2020) Imaging and quantifying homeostatic levels of intracellular silicon in diatoms. Sci Adv 6:eaaz7554

Lewin JC (1955) The capsule of the diatom *Navicula pelliculosa*. J Gen Microbiol 13:162–169

Lind JL, Heimann K, Miller EA, van Vliet C, Hoogenraad NJ, Wetherbee R (1997) Substratum adhesion and gliding in a diatom are mediated by extracellular proteoglycans. Planta 203:213–221

Liu H, Chen M, Zhu F, Harrison PJ (2016) Effect of diatom silica content on copepod grazing, growth and reproduction. Front Mar Sci 3:1–7

Losic D (2017) Diatom nanotechnology: Progress and emerging applications. Royal Society of Chemistry, London

Marron AO, Ratcliffe S, Wheeler GL, Goldstein RE, King N, Not F, de Vargas C, Richter DJ (2016) The evolution of silicon transport in eukaryotes. Mol Biol Evol 33:3226–3248

Martin-Jézéquel V, Hildebrand M, Brzezinski MA (2000) Silicon metabolism in diatoms: implications for growth. J Phycol 36:821–840

Matsunaga S, Sakai R, Jimbo M, Kamiya H (2007) Long-chain polyamines (LCPAs) from marine sponge: possible implication in spicule formation. ChemBioChem 8:1729–1735

Michael A (2011) Molecular machines encoded by bacterially-derived multi-domain gene fusions that potentially synthesize, N-methylate and transfer long chain polyamines in diatoms. FEBS Lett 585:2627–2634

Milligan AJ, Morel FMM (2002) A proton buffering role for silica in diatoms. Science 297:1848–1850

Mock T, Samanta MP, Iverson V, Berthiaume C, Robison M, Holtermann K, Durkin C, Bondurant SS, Richmond K, Rodesch M, Kallas T, Huttlin EL, Cerrina F, Sussman MR, Armbrust EV (2008) Whole-genome expression profiling of the marine diatom *Thalassiosira pseudonana* identifies genes involved in silicon bioprocesses. Proc Natl Acad Sci U S A 105:1579–1584

Mort AJ, Lamport DTA (1977) Anhydrous hydrogen fluoride deglycosylates glycoproteins. Anal Biochem 82:289–309

Motomura K, Ikeda T, Matsuyama S, Abdelhamid MAA, Tanaka T, Ishida T, Hirota R, Kuroda A (2016) The C-terminal zwitterionic sequence of CotB1 is essential for biosilicification of the *Bacillus cereus* spore coat. J Bacteriol 198:276–282

Nakajima T, Volcani BE (1969) 3,4-Dihydroxyproline: a new amino acid in diatom cell walls. Science 164:1400–1401

Nakajima T, Volcani BE (1970) ε-N-trimethyl-L-δ-hydroxylysine phosphate and its nonphosphorylated compound in diatom cell walls. Biochem Biophys Res Commun 39:28–33

Nemoto M, Maeda Y, Muto M, Tanaka M, Yoshino T, Mayama S, Tanaka T (2014) Identification of a frustule-associated protein of the marine pennate diatom *Fistulifera sp.* strain JPCC DA0580. Mar Genomics 16:39–44

Nemoto M, Iwaki S, Moriya H, Monden Y, Tamura T, Inagaki K, Mayama S, Obuse K (2020) Comparative gene analysis focused on silica cell wall formation: identification of diatom-specific SET domain protein methyltransferases. Mar Biotechnol 22:551–563

Otzen D (2012) The role of proteins in biosilicification. Scientifica 2012:867562

Pawolski D, Heintze C, Mey I, Steinem C, Kröger N (2018) Reconstituting the formation of hierarchically porous silica patterns using diatom biomolecules. J Struct Biol 204:64–74

Pickett-Heaps JD, Schmid AMM, Edgar LA (1990) The cell biology of diatom valve formation. In: Round FE, Chapman DJ (eds) Progress in phycological research. Biopress, Bristol, pp 1–168

Poulsen N, Kröger N (2004) Silica morphogenesis by alternative processing of silaffins in the diatom *Thalassiosira pseudonana*. J Biol Chem 279:42993–42999

Poulsen N, Sumper M, Kröger N (2003) Biosilica formation in diatoms: characterization of native silaffin-2 and its role in silica morphogenesis. Proc Natl Acad Sci U S A 100:12075–12080

Poulsen N, Chesley PM, Kröger N (2006) Molecular genetic manipulation of the diatom *Thalassiosira pseudonana* (Bacillariophyceae). J Phycol 42:1059–1065

Poulsen N, Scheffel A, Sheppard VC, Chesley PM, Kröger N (2013) Pentalysine clusters mediate silica targeting of silaffins in *Thalassiosira pseudonana*. J Biol Chem 288:20100–20109

Raven JA (1983) The transport and function of silicon in plants. Biol Rev 58:179–207

Raven JA, Waite AM (2004) The evolution of silicification in diatoms: inescapable sinking and sinking as escape? New Phytol 162:45–61

Reimann BEF, Lewin JC, Volcani BE (1965) Studies on the biochemistry and fine structure of silica shell formation in diatoms. I. The structure of the cell wall of *Cylindrotheca fusiformis* Reimann and Lewin. J Cell Biol 24:39–55

Reimann BEF, Lewin JC, Volcani BE (1966) Studies on the biochemistry and fine structure of silica shell formation in diatoms. II. The structure of the cell wall of *Navicula pelliculosa* (Bréb.) Hilse. J Phycol 2:74–84

Richthammer P, Börmel M, Brunner E, van Pée K-H (2011) Biomineralization in diatoms: the role of silacidins. Chembiochem 12:1362–1366

Scheffel A, Poulsen N, Shian S, Kröger N (2011) Nanopatterned protein microrings from a diatom that direct silica morphogenesis. Proc Natl Acad Sci U S A 108:3175–3180

Sheppard VC, Poulsen N, Kröger N (2010) Characterization of an endoplasmic reticulum-associated silaffin kinase from the diatom *Thalassiosira pseudonana*. J Biol Chem 285:1166–1176

Shimizu K, Del Amo Y, Brzezinski MA, Stucky GD, Morse DE (2001) A novel fluorescent silica tracer for biological silicification studies. Chem Biol 8:1051–1060

Shrestha RP, Tesson B, Norden-Krichmar T, Federowicz S, Hildebrand M, Allen AE (2012) Whole transcriptome analysis of the silicon response of the diatom *Thalassiosira pseudonana*. BMC Genomics 13:499

Spinde K, Kammer M, Freyer K, Ehrlich H, Vournakis JN, Brunner E (2011) Biomimetic silicification of fibrous chitin from diatoms. Chem Mater 23(11):2973–2978

Sumper M (2002) A phase separation model for the nanopatterning of diatom biosilica. Science 295:2430–2433

Sumper M, Brunner E (2008) Silica biomineralisation in diatoms: the model organism *Thalassiosira pseudonana*. ChemBioChem 9:1187–1194

Sumper M, Lehmann G (2006) Silica pattern formation in diatoms: species-specific polyamine biosynthesis. Chembiochem 7:1419

Sumper M, Brunner E, Lehmann G (2005) Biomineralization in diatoms: characterization of novel polyamines associated with silica. FEBS Lett 579:3765

Sumper M, Hett R, Lehmann G, Wenzl S (2007) A code for lysine modifications of a silica biomineralizing silaffin protein. Angew Chem Int Ed 46:8405–8408

Suroy M, Moriceau B, Boutorh J, Goutx M (2014) Fatty acids associated with the frustules of diatoms and their fate during degradation - a case study in *Thalassiosira weissflogii*. Deep-Sea Res I Oceanogr Res Pap 86:21–31

Swift DM, Wheeler AP (1992) Evidence of an organic matrix from diatom silica. J Phycol 28:202–209

Tesson B, Hildebrand M (2013) Characterization and localization of insoluble organic matrices associated with diatom cell walls: insight into their roles during cell wall formation. PLoS One 8:e61675

Tesson B, Masse S, Laurent G, Maquet J, Livage J, Martin-Jézéquel V, Coradin T (2008) Contribution of multi-nuclear solid state NMR to the characterization of the *Thalassiosira pseudonana* diatom cell wall. Anal Bioanal Chem 390:1889–1898

Tesson B, Genet MJ, Fernandez V, Degand S, Rouxhet PG, Martin-Jézéquel V (2009) Surface chemical composition of diatoms. ChemBioChem 10:2011–2024

Tesson B, Lerch SJL, Hildebrand M (2017) Characterization of a new protein family associated with the silica deposition vesicle membrane enables genetic manipulation of diatom silica. Sci Rep 7:1–13

van de Poll WH, Vrieling EG, Gieskes WWC (1999) Location and expression of frustulins in the pennate diatoms *Cylindrotheca fusiformis*, *Navicula pelliculosa* and *Navicula salinarium* (Bacillariophyceae). J Phycol 35:1044–1053

Volcani BE (1981) Cell wall formation in diatoms: morphogenesis and biochemistry. In: Silicon and siliceous structures in biological systems. Springer, New York, pp 157–200

von Stosch HA (1981) Structural and histochemical observations on the organic layers of the diatom cell wall. In: Proceedings of the 6th international symposium on recent and fossil diatoms. O. Koeltz, Königstein, Germany, pp 231–252

Vrieling EG, Gieskes WWC, Beelen TPM (1999) Silicon deposition in diatoms: control by the pH inside the silicon deposition vesicle. J Phycol 35:548–559

Wenzl S, Deutzmann R, Herr R, Hochmuth E, Sumper M (2004) Quarternary ammonium groups in silica-associated proteins. Angew Chem Int Ed 43:5933–5936

Wenzl S, Hett R, Richthammer P, Sumper M (2008) Silacidins: highly acidic phosphopeptides from diatom shells assist in silica precipitation *in vitro*. Angew Chem Int Ed 47:1729–1732

Willis A, Chiovitti A, Dugdale TM, Wetherbee R (2013) Characterization of the extracellular matrix of *Phaeodactylum tricornutum* (Bacillariophyceae): structure, composition, and adhesive characteristics. J Phycol 49:937–949

Wustmann M, Poulsen N, Kröger N, van Pée KH (2020) Chitin synthase localization in the diatom *Thalassiosira pseudonana*. BMC Mater 2:10

Yee DP, Hildebrand M, Tresguerres M (2020) Dynamic subcellular translocation of V-type H^+-ATPase is essential for biomineralization of the diatom silica cell wall. New Phytol 225:2411–2422