Applied Spectroscopy
February 2022, Volume 76 Issue 2 Pages 235-254
https://doi.org/10.1177/00037028211041172
https://archimer.ifremer.fr/doc/00721/83307/

Attenuated Total Reflection (ATR) Micro-Fourier Transform Infrared (Micro-FT-IR) Spectroscopy to Enhance Repeatability and Reproducibility of Spectra Derived from Single Specimen Organic-Walled Dinoflagellate Cysts

Meyvisch Pjotr [1, *], R. Gurdebeke Pieter [1], Vrielinck Henk [2], Mertens Kenneth [3], Versteegh Gerard [4], Louwye Stephen [1]

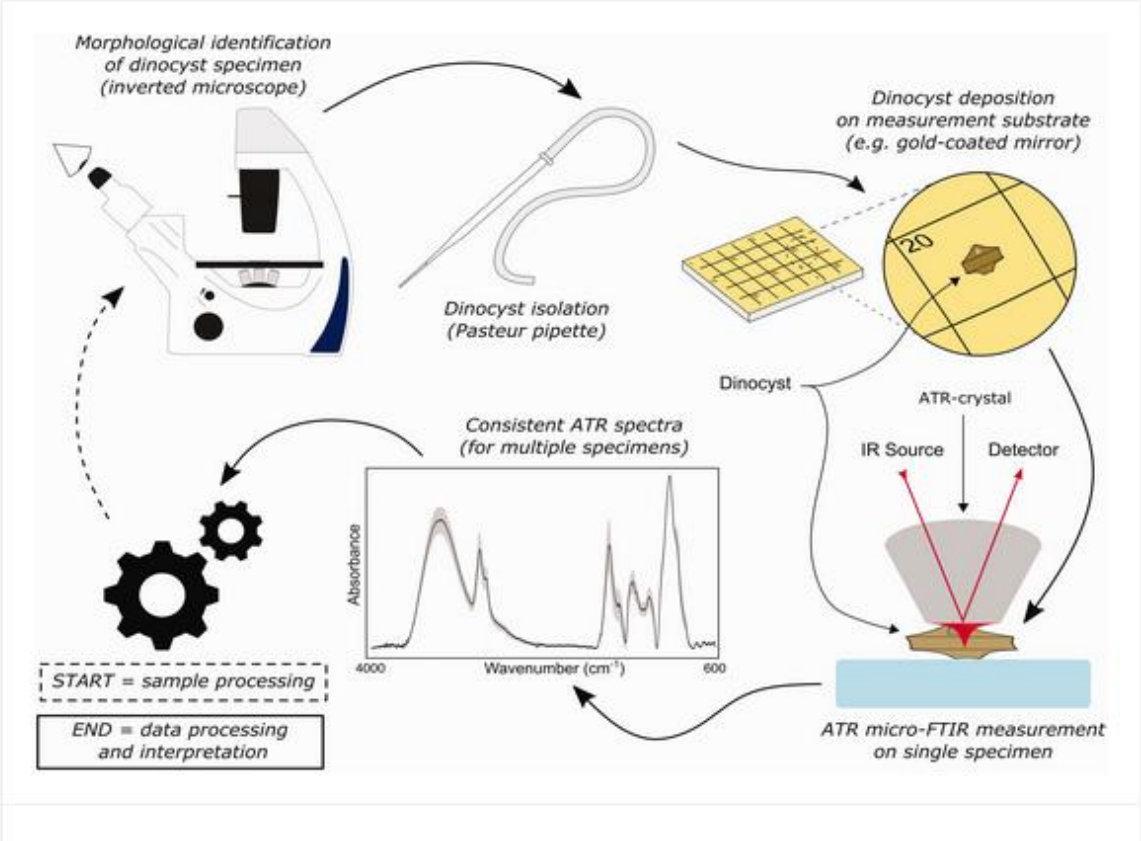
[1] Department of Geology, Ghent University, Ghent, Belgium
[2] Department of Solid-State Sciences, Ghent University, Ghent, Belgium
[3] Ifremer, LITTORAL, Concarneau, France
[4] Marine Biochemistry Group, Alfred-Wegener-Institute, Bremerhaven, Germany

* Corresponding author : Pjotr Meyvisch, email address : pjotr.meyvisch@ugent.be

Abstract :

The chemical composition of recent and fossil organic-walled dinoflagellate cyst walls and its diversity is poorly understood and analyses on single microscopic specimens are rare. A series of infrared spectroscopic experiments resulted in the proposition of a standardized attenuated total reflection micro-Fourier transform infrared-based method that allows the collection of robust data sets consisting of spectra from individual dinocysts. These data sets are largely devoid of nonchemical artifacts inherent to other infrared spectrochemical methods, which have typically been used to study similar specimens in the past. The influence of sample preparation, specimen morphology and size and spectral data processing steps is also assessed within this methodological framework. As a result, several guidelines are proposed which facilitate the collection and qualitative interpretation of highly reproducible and repeatable spectrochemical data. These, in turn, pave the way for a systematic exploration of dinocyst chemistry and its assessment as a chemotaxonomical tool or proxy.

Graphical abstract



Keywords : Dinocysts, dinoflagellates, attenuated total reflection, ATR, Fourier transform infrared spectroscopy, FT-IR, micropaleontology, chemotaxonomy, chemometrics, methodology

1. Introduction

Dinoflagellates are mostly unicellular, eukaryotic protists that occur in a wide array of aquatic environments. Most dinoflagellates are planktonic, although benthics, symbionts and parasites also exist. They are either photosynthesizing primary producers (autotrophs), consumers of other micro-organisms (heterotrophs) or a combination of both (mixotrophs). During the sexual phase of their complex life cycle, about 13–16% of living dinoflagellate species produce a non-motile resting cyst (dinocyst), usually 20–100 μm in diameter.[1] These potentially highly resistant dinocysts sink and can preserve in sediments and in some cases have been shown to remain alive for more than a century.[2] In environmental conditions which favor vegetative growth, dinocysts can germinate and motile cells reinoculate the water column. Most

commonly dinocyst are organic-walled with a composition remarkably different from that of the motile stage, but calcareous and silicious forms exist as well. The chemical resistance of the dinocyst wall procures a potential for fossilization, and fossil dinocysts have been found dating as far back as the Triassic, over 200 million years ago.

Many studies on dinocysts focus on morphological (e.g., body shape, dimensions and ornamentations) characteristics which are used in taxonomical, biostratigraphical, (paleo)ecological, (paleo)climatological and palaeoceanographical applications.[3] Biological studies often include molecular data for investigating the taxonomy and complex life cycles of certain toxin-producing dinoflagellates, which can cause large harmful algal blooms in modern-day aquatic environments.[4] In contrast, the chemical compositions of organic-walled dinocyst walls are largely unexplored. The microscopic cyst sizes and the possibility of chemical contaminations can impose difficulties on their chemical characterization. Despite this, the applicability of dinocyst chemical data can be considered large and serves multiple purposes. It allows to better understand the mechanisms behind dinocyst taphonomy, their compositional makeup and evolution through time, as well as to evaluate the potential of such data as a chemotaxonomical tool and proxy. Traditionally, the walls of organic-walled dinoflagellate resting cysts are considered to consist of a suite of highly resistant biomacromolecules called 'dinosporin'.[5]

Past studies on organic-walled dinocyst wall chemistry mainly applied methods based on chromatography and mass spectrometry in conjunction with hydrolysis, pyrolysis and thermally assisted hemolysis and methylation[6–10], as well as Fourier Transform Infrared (FTIR) spectroscopy on individual cysts by focusing and collecting the infrared (IR) beam with microscope objectives (micro-FTIR).[6,8–14] These studies have already indicated that dinosporin differs from similar resistant biopolymers like sporopollenin as found in pollen, and algeanan as found in green algae, and is likely to be a highly crosslinked (β-1,4) carbohydrate-based biopolymer. Furthermore, micro-FTIR data demonstrates a large (geo)chemical dinosporin variability in recent and fossil dinocysts, which seems to be linked to dinoflagellate species feeding strategies as well as phylogeny, environment and preservation.[15]

The chemical information retrieved in FTIR spectroscopy [16,17] originates from the excitation of IR-active vibrations of the molecules in a sample through the absorption of IR light. The absorption at a specific IR frequency, typically expressed in wavenumbers (= $\bar{\nu}$ = 1/λ, with λ = wavelength in cm), is determined by the vibrational spectrum of the molecules in the sample and their amount in the region probed by the IR beam. With respect to dinocysts, vibrational absorption spectra typically consist of relatively broad ($\geq$ 30 cm^{-1}) but specific bands (i.e., molecular group absorption bands) in the mid-IR region (4000–400 cm^{-1}). Transmission, reflection and attenuated total reflection (ATR) are considered the three major IR-spectroscopic sampling modes[18,19], and differ in the relative positions of the IR source and detector with respect to the sample and the way the IR light is guided through it (**Figure 1**). For organic-walled dinocysts measured in reflection mode part of the IR light passes through the sample before being reflected to the detector. Therefore, we further call this sampling mode transflection.

The present study is focused on the further development, assessment, applicability, and outlook of micro-FTIR spectroscopy as a tool for the (geo)chemical characterization of single, organic-walled dinocyst specimens. The most convenient

way to retrieve clean dinocysts (without cell contents and attached external debris) from a sample while maintaining taxonomic control is by individual manipulation by hand (**Figure 2**). FTIR spectroscopy of macroscopic quantities of dinocysts, e.g., pressed in KBr pellets or directly onto the ATR prism of a bench spectrometer, is very time consuming and difficult with respect to maintaining sample purity and is not considered. Here, a method based on ATR (hence ATR micro-FTIR) spectroscopy on single, organic-walled dinocysts is presented and the recorded spectra are compared to spectra collected via other sampling modes (transmission and transflection), which were typically employed in previous micro-FTIR studies on these resting stages. Building on previous efforts for erecting a standardized protocol for measuring dinocyst via micro-FTIR[20], the present study captures and highlights a range of methodological aspects which prevail during sample preparation, data collection, processing and interpretation. Finally, future outlooks for the use of micro-FTIR spectroscopy on single, organic-walled dinocyst specimens are discussed, as well as its use in conjunction with complementary methods. Major standing questions in this field of research are outlined.

2. Materials and methods

2.1 Sample preparation

Modern surface sediments and a live culture containing dinocysts from different nearshore marine or lagoonal settings around the world (**Table 1, Figure S1**) were prepared for micro-FTIR analyses via partly different series of physical and chemical processing steps (**Table 1**). A brief description and explanation of these steps is given below:

- *Palynological maceration* – to remove inorganic compounds, sediments were treated with the following strong acids (in order): 7.3% HCl (24h at room temperature, to dissolve carbonate particles), 40% HF (48h at 60°C, to dissolve silicates) and 2N HCl (cycles of 24h at 60°C, until the mineral fraction was fully dissolved). The remaining organic matter residue was neutralized by iteratively adding and decanting distilled water. The final residue was stored in glass vials. Palynological macerations are routinely used in studies on fossil dinocysts liberating them from their inorganic rock matrix, and different methods exist. For further details see a closely similar protocol.[21] Only some sediment samples in this study have undergone palynological maceration (**Table 1**) in order to compare the spectral outcomes of different sample preparation procedures.
- *Ultrasonication* – samples were dispersed in distilled water in a glass beaker and afterwards immersed in an ultrasound water bath (Branson 2510 Ultrasonic Cleaner at 40 kHz) to achieve two goals: (i) to break up non-germinated dinocysts, effectively removing their cell contents; (ii) to remove contaminating particles of variable sizes sticking to the outside of the cyst walls and processes. The number of consecutive sonication steps varied depending on the visual cleanliness of the dinocysts after a fixed duration of 30s and was carried out prior to filtering or density separation.

- *Filtering* – since most extant dinocyst species rarely exceed 100 µm in diameter, sediments and residues were first filtered on a steel 150 µm mesh and caught on 5, 8 or 20 µm nylon MF-Millipore™ MCE membranes, using ultrapure (Millipore Milli-Q®) or distilled water. Subsequent filtering on nylon membranes was preceded by ultrasonication to ensure maximum disintegration of aggregating particles.
- *Rinsing with organic solvents* – in order to remove (a)polar compounds, often lipids, sticking to dinocysts some samples were extracted for 30 min, while shaking with a 1:1 mixture of methanol and dichloromethane (MeOH:DCM) in a sealed plastic centrifuge tube. The residue was subsequently caught on a 8 µm nylon filter and rinsed three times with ultrapure water.[11,22–24]
- *Density separation* – for some samples dense, non-dinocyst particles were removed via the density separation method of Bolch[25] that uses an aqueous solution of non-toxic sodium polytungstate (SPT), combined with centrifugation. A higher SPT density of 2.1 g/cm^3 was used, in comparison to Bolch's 1.4 g/cm^3, which works well for both recent[26] and fossil dinocyst samples.[27]
- *Fixation* – some final residues were fixed by adding a few drops of ethanol (EtOH) to prevent bacterial growth which could chemically contaminate dinocysts.

From the final processed residues, single dinocysts were isolated for IR measurement (**Figure 2**). Droplets were examined under a Zeiss Primovert inverted microscope (100x to 400x magnification). Visually clean dinocysts were identified and manipulated using a drawn glass Pasteur pipette with an attached rubber suction tube. Individual specimens were iteratively cleaned by depositing them in droplets of distilled water, rinsing the Pasteur pipette with distilled water in between each iteration. Afterwards, the specimens were carefully deposited onto the micro-FTIR measurement substrate, making sure no visual drying spot was cast around their periphery. This was ensured either by precise control of the pressure within the suction tube, resulting in a small and rapidly evaporating water droplet containing the dinocyst, or by sucking the water drop away underneath the deposited dinocyst by using another Pasteur pipette with an aperture diameter smaller than that of the respective cyst. Specimens measured in transmission mode were deposited on either a 7.7 mm thick and 2.5 cm wide, circular ZnSe-window (~70% IR transparency) or a 0.5 mm thick 2.5 by 2.5 cm, rectangular Si-window (~54% IR transparency from 4000–1550 cm^{-1}, with well-known absorption bands from 1550–600 cm^{-1} which hardly influence spectra resulting from thin samples, which is the case here). ZnSe and Si were chosen as they are both partially transmitting as well as reflecting IR light in the targeted spectral range (4000–600 cm^{-1}), allowing to (re)measure the exact same specimens in all three sampling modes. CaF_2 was not considered in this respect, given its strong absorption below 1000 cm^{-1}. Samples which were not measured in transmission (only transflection and ATR) were deposited on a 5 by 5 cm Mid-Infrared Enhanced Au-mirror (Thorlabs®), as this substrate was routinely used in the past for dinocyst transflection measurements. Specimens were measured within 24h after sample isolation/deposition.

2.2 Micro-FTIR measurements and data analysis

A total of six, morphologically distinct dinocyst taxa have been analyzed via different micro-FTIR spectroscopy sampling modes, with varying antecedent chemical treatments and on different measurement substrates (**Table 2**). For each taxon, different spectra were recorded on multiple (2–10) empty (i.e., without cell contents) and visually clean individual specimens, serving as biological replicates. Unique specimens belonging to a given species, location and measurement substrate have been analyzed in a nondestructive sampling mode (i.e., transmission and/or transflection) prior to ATR measurement. In those cases, this allows for a one-on-one comparison of spectra from the same specimen, as recorded via different micro-FTIR methods. ATR is considered to be a destructive technique with respect to dinocyst samples, since they are naturally hollow structures which become deformed (flattened) or sometimes break up into smaller fragments by the pressure of the ATR crystal. Furthermore, the cysts often stick to the crystal after measurement and are lost when cleaning the crystal.

IR spectra from the Xiamen Bay and Patricia Bay localities as well as the *Lingulodinium machaerophorum* culture samples measured on an Au-mirror were recorded with a Bruker Invenio-S spectrometer with CenterGlow IR source coupled to a Hyperion 1000 IR microscope located at the Marum Research Faculty (Bremen University, Germany), while others were recorded with a Bruker Vertex 80v spectrometer with Globar IR source coupled to a Hyperion 2000 IR microscope located at the Department of Solid State Sciences (Ghent University, Belgium). In both setups, infrared spectral ranges of 4500–600 cm^{-1} were recorded using a liquid N_2-cooled mercury cadmium telluride (MCT) detector in the microscope and KBr beam splitter in the bench spectrometer. The Akkeshi Bay and *Votadinium calvum* spectra from Patricia Bay were collected with a 2 cm^{-1} resolution and averaged over 100 scans, while the others are collected with a 4 cm^{-1} resolution and averaged over 256 scans. This as part of different standard lab procedures and to test if it makes any significant difference in terms of the overall spectral outcome. Prior to each measurement, the respective specimen was photographed using the built-in camera in the Hyperion microscope unit. This served as an extra visual identification of the measured specimen. Transmission and transflection spectra were recorded by using a x15 (0.4 NA) microscope objective (and a x15 = 0.4 NA condenser in transmission) and by first adapting the rectangular IR aperture size to roughly fit the dimensions of the specimen, then recording a background spectrum on an empty and visually clean area of the measurement substrate and afterwards recording the specimen itself. For sample measurement, the IR beam was focused on the top of the measurement substrate. Prior to transmission measurements the condenser was aligned and focused for maximum signal intensity. ATR spectra were obtained by using a germanium Internal Reflection Element (IRE) with a tip diameter of 100 µm mounted on the x20 (0.6 NA) microscope objective of the Hyperion microscope. Before the first and after each subsequent sample measurement the IRE tip was cleaned with isopropanol alcohol and a cotton tissue. A background spectrum was recorded every five measurements, or earlier when previously recorded spectra contained large atmospheric contributions due to changing atmospheric conditions between background and sample measurement. The background spectra were taken by placing a supportive plastic ring on the stage around the ATR crystal and the measurement substrate, applying

pressure up to pressure level 1 (~0.5 N)[28] and measuring the IR spectrum of the atmosphere surrounding the IRE tip. The ATR micro-FTIR sample spectra were recorded by lowering the IRE onto the specimen, making contact until pressure level 1 was reached.

For visualization, comparison and interpretation purposes, subsequent spectral processing of the raw data was carried out in OPUS 8.2.21 (Bruker) and included in the following order: atmospheric compensation (removal of residual gas phase CO_2 and H_2O absorption traces in the spectra), Savitzky-Golay (smoothing) filtering (second order polynomial and a window size of 13)[29,30], baseline correction (rubberband correction using polynomes, 128 baseline points and one to several iteration(s) until the onset of the absorption band near 3700 cm^{-1} and the onset of the fingerprint region (1750–600 cm^{-1}) near 1800 cm^{-1} reached approximately zero absorbance values) and min-max normalization (plots showing a close-up of the fingerprint region have been normalized to the values within this region). Absorption band groups were assigned following Colthup et al.[31], Coates[32], Bogus et al.[12] and Jardine et al.[33]. All spectra and additional statistics were plotted using RStudio.[34]

For this study, the quality of spectra produced by these different approaches is mainly assessed by means of the measurement repeatability (i.e., the similarity of results obtained from the same dinocyst species using the same measurement procedure, numerically expressed by the average standard deviations from the mean of the biological replicates), reproducibility (i.e., the similarity between the measurement results from the same dinocyst species obtained at two independent labs) and signal-to-noise ratio (SNR) and an explanation for the spectral variations is presented. Spectra are considered to show high consistency when they exhibit relatively high repeatability, reproducibility and SNRs.

3. Results and discussion

The measured micro-FTIR spectra (**Figures 3 and 4)** show two main spectral regions with absorption bands: (i) the high wavenumber region between 3700–2700 cm^{-1} with (relatively) broad and strong absorption bands and (ii) a fingerprint region (1750–600 cm^{-1}) which can roughly be divided into three subregions with characteristic molecular group absorption bands which are often seen in dinocyst micro-FTIR spectra.[9–14] An overview on common dinocyst group frequencies, their origins and assignments is given in **Table 3**. Differences between the three used sampling modes are evaluated in section 3.1, which demonstrates that ATR FTIR is the most reliable method. Further effects of sample processing and data processing on ATR FTIR spectra are discussed in section 3.2.

3.1 Comparison of sampling modes

3.1.1. What method is most consistent?

Spectra of empty, cultured *Lingulodinium machaerophorum* cysts, have been recorded via all three sampling modes and on two different IR-transparent substrates (**Table 2**, **Figure 3, Figure S2)**. Since the setup geometries and the pathways and

interactions of IR light differ between the sampling modes and measurement substrates, their resulting spectra can differ significantly in terms of shape, intensity, and SNR, even when measured on the same specimen. For the analysis of single cysts, each of the FTIR sampling modes has its advantages and limitations (**Table 4**) which to a large degree can be used to explain the observed spectral variability. The measurement repeatability and the reproducibility are two metrics which can be used to assess the quality and reliability of collected data. The repeatability can be quantified by calculating, for each recorded wavenumber, the average standard deviation from the mean (i.e., the averaged spectrum of the biological replicates). For *L. machaerophorum* culture data, the spectra collected in ATR systematically show the highest repeatability over those collected in transmission and transflection, independent of the measurement substrate type (**Figure 3**). Additional to the cultured samples, a total of four *L. machaerophorum* cysts extracted from sediment (Diana Lagoon, **Table 2**) were measured in ATR on an Au-mirror substrate. The fingerprint regions of these spectra (dotted lines in **Figure 3c**) show high similarities in terms of absorption band positions and shapes with those from the cultured spectra. Given that the culture and sediment datasets were collected using different instruments located in independent labs (the culture-derived spectra at MARUM, the sediment-derived spectra at Ghent University), this also suggests a high reproducibility for ATR spectra. Differences in relative absorption band heights in the fingerprint between culture and sediment *L. machaerophorum* cysts might be attributed to environmental factors, chemical changes which occurred during culturing or indiscernible (extra)cellular materials sticking to the cultured cysts. Large spectral differences can be detected between the transmission and transflection spectra recorded on different measurement substrates. For transmission spectra specifically, high frequency background noise can be seen due to low signal intensities (= low SNR), which assumably for a large part causes the low reproducibility. In case of the of the ZnSe-window, the low SNR are likely due to surface scratches, which are visible on some visual images and cause scattering of IR light. We expect no significant contribution of the window's thickness (7.7 mm) to the low SNR's, nor from cropping of IR light near the edges since all specimens were deposited in a small 3 by 3 mm cluster in the center of the 2.5 cm wide window and given the 1.2 mm spot size of the 0.4 NA microscope objective. Furthermore, this particular window showed no chromatic aberrations on visual images or aberrations to plane parallelism. Transflection spectra generally show the lowest repeatability within but also between measurement substrates, which can be explained by the differences in reflectance between the substrates (from ZnSe being only about 30% reflecting and the Au-mirror nearly 100% in IR) and the associated complex scattering effects (see following paragraph). In contrast, ATR spectra (after data processing) show high consistency suggesting a very minimal to negligible lasting influence of the measurement substrate (**Figure S3**).

From our experimental results and the advantages and limitations listed in **Table 4** it appears that, apart from being nondestructive and able to measure the sample over its entire thickness, transmission and transflection measurements offer few advantages over ATR measurements for the analysis of single specimen dinocysts. The main reason ATR with a high refractive IRE outperforms the other two sampling modes is that in the latter, (electromagnetic) scattering effects occur when the beam

crosses a sharp refractive index (n) boundary.[35] These scattering effects cause spectral distortions in terms of absorption band intensities, band positions, relative band proportions and baseline effects and thus obscure the true absorbance spectrum.[36,37,46–50,38–45] For some sample setups and geometries, these spectral distortions can be partly quantified and corrected via model-based correction algorithms which often need pure absorbance spectra as a reference.[37,51–59] In the past few years deep learning approaches have made their advance in FTIR spectroscopy.[60–63] Worth noticing is that recently developed deep neural network architectures can successfully approximate complex, often computationally heavy, model-based correction algorithms.[64] The currently available correction algorithms are often optimized for transmission spectra, as the distortions in transflection spectra are in general more severe. In the latter method the scattering effects are essentially doubled, making correcting such spectra practically impossible. This, together with the irregular and unique shapes of individual dinocysts (morphological variability between and within species, but also differential collapse during desiccation on the substrate), explains the low repeatability and large differences between transflection spectra on different substrates (**Figure 3**). Therefore, for single dinocysts, transmission is evaluated as the most reliable nondestructive technique to use, provided that the sample is thick and large enough and the measurement substrate transparent enough to create sufficient signal intensities, as well as smooth enough to limit scattering effects.

Scattering effects are absent in ATR measurements of *L. machaerophorum* since the internally reflected IR beam does not leave the IRE (i.e., does not cross a refractive index boundary, **Figure 1**), resulting in an increased reproducibility (**Figure 3**).[65] The evanescent wave created at the tip of the IRE propagates into the sample in close vicinity of the tip, though only very shallow. For a Ge IRE (n = 4.0) at about 30°, which is used in this study, measuring dinocysts ($\mathrm{n} \approx 1.50$) the penetration depth ranges from about 0.15 µm (at 4000 cm^{-1}) to about 1 µm (at 600 cm^{-1}). Organic walls of dinocyst of modern species generally range from one to several µm in thickness. This implies that, especially at higher wavenumbers, possibly only the outermost layer(s) of the compressed cyst wall is/are probed via ATR micro-FTIR and that the absorption bands in the fingerprint region will be relatively stronger compared to those in the spectral region between 3700–2700 cm^{-1}. Though free of scattering effects, ATR is – to a lesser degree – also prone to physical artifacts caused by a number of factors: (i) inhomogeneities in the refractive indices of sample substructures; (ii) quality of the sample contact with the IRE and (iii) depth of penetration.[66,67] Incomplete sample contact and/or small sample thicknesses can also give rise to fringing effects in ATR spectra, possibly due to standing waves, but these were not observed in this study. Due to the high repeatability within and between ATR spectra recorded on different measurement substrates, our experimental results suggest that the physical artifacts in ATR spectra of dinocysts measured in this study are small, in particular when compared to the scattering distortions observed in transmission and transflection spectra. When present, physical artifacts presumably cause mostly minor changes in absorption band intensities, rather than shifts in band positions which can be considered much more problematic when it comes to qualitative molecular identification. Furthermore, absorption band intensity changes in spectra can to a

certain degree be corrected via model-based and deep learning algorithms.[59,61] The near absence of high-frequency noise in the ATR spectra also implies that the measured specimens were sufficiently large and a good contact with the IRE tip was established. Although the diameter of the circular IRE tip is fixed at 100 µm, the typical probing area (i.e., spatial resolution) is smaller depending on the refractive index of the crystal, the shape of the aperture and the diffraction limit of the microscope objective.[68,69] The Ge ATR objective used in this study is reported to be able to measure samples as small as the order of magnitude of 5 µm. From our own experiences, ATR spectra from very small dinocyst specimens (< 30 × 30 µm when compressed) tend to become noisy and this is seen as a lower size limit. For all the above-mentioned reasons, micro-ATR is evaluated as the most reliable sampling mode for collecting spectra of single specimen dinocysts. Since this evaluation is mostly made on the basis a repeatability metric, it highlights the importance of collecting multiple technical and biological replicates when applying micro-FTIR spectroscopy to microscopic bioparticles. Other micro-FTIR studies on micro-organisms generally collect a minimum of three replicates.[70–72] Our experiences learn that, in general, there is some intraspecific variation in dinocyst spectra from the same sample, but this tends to be small if the cysts are clean. Usually, 3–5 spectra on different cysts from the same species are sufficient to capture most of this variation. We have tested this for several species with > 20 unique spectra or specimens (not presented in this study).

3.1.2. Causes of higher scattering in transflection: effects of morphology and size on scattering

Transflection and ATR micro-FTIR spectra from additional dinocyst taxa are compared (**Figure 4**), mainly to highlight the intense scattering artifacts in the transflection spectra, their link with the sample morphologies and to illustrate that for all taxa ATR spectra prove to be much more consistent and devoid of scattering artifacts. Furthermore, a one-to-one comparison of transflection and ATR spectra from some single specimens shows that spectra collected in transflection mode, which was often applied in past micro-FTIR dinocyst studies[12–14,24,73], can contain major distortions in terms of (relative) absorption band heights and band positions, which significantly hampers their further qualitative and quantitative interpretation.

Transflection spectra from *Brigantedinium* spp. contain major Mie-type scattering artifacts[54,55], clearly seen as sinusoidal intensity variations between 2800–1800 cm^{-1} and to a lesser degree in the fingerprint region. This is not surprising given that the morphology of small (typically 30–50 µm), empty *Brigantedinium* cysts closely resembles that of a near perfect hollow sphere (**Figure 4**) similar to spheroidal pollen grains which produce strong Mie scattering.[46] To a lesser degree, sinusoidal scattering artifacts can be observed in transflection spectra of *Quinquecuspis concreta* and *Spiniferites bentorii*. *Q. concreta* is relatively large (typically 60–80 µm), smooth and pentagonal-shaped. *S. bentorii* is only a little bit smaller and oval, but with relatively short (15–25 µm), distally trifurcating processes (**Figure 4**). These are both shapes which differ significantly from that of an ideal sphere, and they produce complex, irregular scattering effects which to a certain degree obscure the regular sinusoidal

patterns as observed in *Brigantedinium* spectra. Cysts of *Trinovantedinium applanatum* and *Votadinium calvum* are typically relatively large (70–100 µm), slightly ornamented (i.e., low topography processes) to smooth and dorsoventrally flattened (**Figure 4**). It is clear from the presented transflection spectra that these morphologies produce less intense scattering artifacts: the spectra show a higher repeatability and more closely resemble the scatter-free ATR spectra. Overall, it is clear that dinocyst morphology as well as size influence the quality of spectra recorded in transflection mode, with smaller, spherical and more ornamented forms leading to less consistent, scatter-distorted spectra. Our experiments prove that neither dinocyst morphology nor size (i.e., above the 30 µm size threshold for producing sufficient signal strengths) significantly reduce the quality of the ATR spectra as they show high consistency throughout. It should be noticed that for very small dinocysts (< 30 µm) multiple specimens can be clustered together on the measurement substrate to achieve sufficient signal strengths. When single-specimen spectra of small cysts are required, transmission measurements using high intensity radiation (e.g., synchrotron) focused near the spatial resolution limit of IR light can be considered the best alternative.

3.2. Other variables important for interpretation of ATR FTIR spectra.

Next to variability inherent to the sample method used, there are effects of sample processing and data processing. These are dealt with in sections 3.2.1 and 3.2.2.

3.2.1. General effects of sample processing on ATR FTIR spectra.

Not all measured specimens of *S. bentorii*, *T. applanatum* and *V. calvum* have undergone the same sample preparation procedure (**Tables 1 and 2**), which could explain some of the observed spectral variability for all sampling modes (**Figure 4**). Notably the use of strong acids like HCl and HF during palynological maceration might alter the molecular structures of dinocyst walls. Acid treatment may turn organic salts into acids, which may redistribute charges in the macromolecule and change the vibrational spectra. Furthermore, acid treatment may hydrolyze part of the macromolecule, e.g., by breaking amide, ester and glycosidic bonds. The loss of organic acids and alcohols may thus be expected. The extent to which this happens depends on the kind of acid used, pH, temperature and duration of the treatment. At room temperature, and with 7.3% HCl the effects are most probably minor.[74] The impact of HF most probably will be removal of minerals sorbed or sticking to the cysts.[75] The concentrations of the chemicals used in this study as well as the reaction temperatures are considered to be too low to induce Maillard reactions, which produce insoluble polymers. The fingerprint regions of the ATR spectra of *T. applanatum* from Olhão Port (no palynological maceration) and Xiamen Bay (palynological maceration) show only minor differences in terms of band positions and intensities, which could possibly reflect environmental factors, intraspecific chemical variations and/or (extra)cellular materials sticking to the cysts. Acid hydrolysis is likely the cause for the loss of free OH groups, which is reflected by a significant intensity reduction in the OH-band (3700–3100 cm^{-1}, **Figure 4**). Transflection spectra of *V. calvum* from Patricia Bay (no palynological maceration) show large absorption bands between 1150–900 cm^{-1}

(typically aromatic or sugar ring absorptions) which are absent in specimens from Xiamen Bay treated with strong acids. Even though these spectra are recorded in transflection mode, the amount of scattering distortions they generally contain appear relatively small. This can be explained by *V. calvum* being relatively large (70–100 µm) and flattened in shape. Therefore, it is safe to assume the additional absorption bands are of chemical origin (presumably due to extracellular weakly bound polysaccharides), rather than scattering related. In this case, acid hydrolysis seems to remove (a) molecular component(s) rich in C-O. One option is that the cyst wall of *V. calvum* consists of multiple layers, one of which is resistant to hydrolysis. Furthermore, similar trends are observed in other *Votadinium* species as well as other recent, brown-colored heterotrophic dinocysts (unpublished data). Though currently still highly understudied, a differential impact of the use of strong acids on the wall composition of recent, and possibly fossil, dinocysts might exist and care should be taken when comparing spectra from samples which underwent a different preparation procedure.

This observation also raises the question whether a standard sample preparation protocol with respect to micro-FTIR spectroscopy on dinocysts should be developed to exclude extrinsic variables as much as possible. Further systematic micro-FTIR spectroscopy experiments on the differential chemical resistances of dinocysts, possibly in conjunction with other (geo)chemical methods (e.g., chromatography and mass spectrometry in combination with hydrolysis and pyrolysis), would allow to formulate a more substantiated answer to this question. As unprocessed fossil dinocysts are often enclosed in consolidated sediments, palynological maceration is indispensable to release the cysts for further manipulation and isolation. In some cases, unconsolidated sediments are also treated with strong acids to dissolve inorganic fractions and to concentrate dinocysts. In most cases, heavy liquid density separation using SPT works just as well for sufficiently concentrating cysts and facilitate further preparations for FTIR measurements without possibly altering the cyst wall chemistries. Usually, no density separation and no palynological maceration is necessary for culture and plankton net samples. The variety in dinocyst sample types will probably prohibit the erection of a universal preparation protocol. Therefore, a standardized sample type-specific preparation procedure is proposed (**Figure 5**). If after the visual inspection step a sample contains large amounts of flocculating amorphous organic matter (AOM) particles, it is nearly impossible to find and isolate visually clean individual dinocysts. An additional processing step can be undertaken to solve this and involves shaking and decanting cycles in ultrapure or distilled water. A study by Anderson et al.[76] reports settling velocities of three different dinocyst species between 0.008–0.013 cm/s in filtered seawater (salinity = 32‰, 22°C). This is equivalent to a downward motion of at least 10 cm in 20 minutes. Ultrapure and distilled water have a lower density at room temperature (~1.00 g/cm^3) than the Percoll-sorbitol-seawater used in the study (1.15 g/cm^3) and the average density of seawater (1.027 g/cm^3).[77] Thus, dinocysts will sink faster in the freshwater media. This means that, when working with a small beaker (ca. 10 cm in height) and following shaking or ultrasonication, dinocysts will no longer be in suspension after settling for about 20 minutes. The materials that are still in suspension after this time span are mostly smaller AOM particles and can thus be safely decanted for an overall cleaner residue. The process of re-suspension, ultrasonication and decanting can be repeated

iteratively if needed. From our experiences, the cysts are usually not contaminated with lipids so that extraction with organic solvents usually has little effect. However, for cysts from organic-rich samples and source rocks, the extraction step is highly recommended.

3.2.2. Options for further handling ATR FTIR spectral datasets

Depending on the aims of the study, FTIR data can be further analyzed and interpreted in different ways ranging from basic spectral visualization and qualitative profiling to exploratory multivariate methods (e.g., Principal Component Analysis, PCA), clustering methods (e.g., Hierarchical Cluster Analysis, HCA, k-means clustering, Fuzzy C-means clustering and density-based clustering, DBSCAN) and classification approaches (e.g., artificial neural networks, ANNs, convolutional neural networks, CNNs, random forest classifiers, RF and deep learning approaches). Based on the results presented here and by comparison with approaches in other micro-FTIR studies, a flowchart is presented which outlines several essential and optional data preprocessing steps prior to specific data analysis methods (**Figure 6**). These steps include:

- *Atmospheric compensation* – residual water vapor and carbon dioxide absorptions are always present in spectra recorded in atmospheric conditions and can be removed by subtracting pure spectra of these gas compounds from the sample spectrum.[78]
- *Spectral quality test* – either based on predefined thresholds (e.g., specific absorbance values, SNRs, intensities of IR water vapor bands and optical fringes)[79] or model-based correction algorithms[80], quality tests can be used to detect and remove outliers, allowing for more sensible further analyses.
- *Savitzky-Golay (SG) filtering and smoothing* – the SG algorithm[29] is most commonly used in FTIR spectroscopy[79] and is able to filter out instrument-derived high frequency noise, typically present in unprocessed spectra. During SG filtering the spectrum is approximated by polynomial least-squares fitting in a moving window, by predefining the polynomial order and window size. Using a second order polynomial and window sizes of 9–25 (dataset-specific) generally results in the best classification models for experimental and simulated mid-infrared (4000–400 cm^{-1}) FTIR data.[30] Setting the window size is often a tradeoff between noise removal and preventing loss of spectral information.
- *Derivation (differentiation) and decomposition* – derivative filters can be used to enhance the resolution of IR spectra, since they are able to resolve overlapping bands as well as reduce or eliminate baseline offsets and slopes. This in turn reduces spectral complexity.[30,79,81–84] and can result in better classification models afterwards.[72,85–93] Since derivation enhances the noise level of spectra, it is only advisable to use on high SNR data.
- *Baseline correction* – where (linear and undulating) baseline effects in transmission and transflection spectra occur due to scattering effects, the presence of a baseline in raw ATR data (**Figure S3**) is mainly due to substrate effects for non-infinite samples (i.e., the penetration depth is larger than the sample thickness). Since the

baseline correction is not able to correct for more complex scattering and physical artifacts, it should be used with care before multivariate analyses, clustering, or classification tasks. This is usually not a large concern when dealing with scatter free ATR spectra measured on Au and ZnSe substrates, where the baseline can be well approximated by a simple exponential function.

- *Normalization* – this can be used to minimize the effects of differing optical pathlengths (e.g., due to varying thickness of the sample), intensity variations in the IR source[79] or scattering (in transmission and transflection) or surface contact variations (ATR), making spectra more intercomparable. Numerous normalization methods exist.[79]
- *(Extended) ATR correction* – this correction is necessary when the goal is to compare ATR data with transmission data and arises from the fact that the penetration depth of the incident IR light in ATR depends on its wavelength, as well as its incidence angle and the refractive indices of the IRE and the sample. Anomalous dispersion causes strong fluctuations in the refractive index of the sample around intense absorption bands.[94] ATR correction is only sensible when the sample thickness is larger than the penetration depth, which is not always the case for dinocysts and should therefore be used with care. As long as the sample is thick enough and if all the spectra in the dataset are collected with the same IRE crystal and angle of incidence, the penetration depth variations will be nearly identical between all samples, since most organic materials have a refractive index of around 1.5.[95] ATR correction is in practice only necessary when comparing measured dinocyst spectra with library reference spectra on pure compounds, which are often measured in transmission mode.
- *Model- or deep learning-based correction algorithms* – these include algorithms designed to correct for spectral distortions caused by nonchemical factors. A popular model-based algorithm is the extended multiplicative signal correction (EMSC), which uses a reference spectrum and polynomial curve fitting.[51,96] With respect to FTIR spectra of biological specimens, a polynomial order of two (EMSC including a quadratic term) is mostly used, since it approximates the relatively broad absorption bands present in these spectra. Extensions to basic EMSC exist to account for specific types of light scattering like resonant Mie-type scattering (RMieS-EMSC) that dominantly occur in spherical scatterers[54,55,59] or scattering in fibers and cylindrical domains.[97] It is however still unclear if EMSC-based algorithms are able to correct perfectly and extract pure absorbance spectra, a feat which might only be currently achievable by other modelling and deep learning approaches,[58,60–62,64] which are nonlinear by nature.[98] For some deep learning classification algorithms the best classification results are observed without any prior data processing, except for assignment of a target class and splitting of the dataset into training, validation and testing sets.[60]

For clarity some of the above-mentioned preprocessing steps have been carried out on and visualized for the ATR data from cultured *L. machaerophorum* cysts, measured on different substrates (**Figure S3**). This figure shows that further processing via PCA, HCA and RF classification allows to distinguish spectra collected on different substrates. Similar workflows using dimensionality reduction, clustering

and classification techniques (e.g., with taxonomic class as the target variable; i.e., chemotaxonomy) can be applied to dinocyst datasets in order to quantify or highlight spectral (dis)similarities and are the scope of future research. Most of these calculations are embedded in software packages which are also used to collect spectra, of which an overview is presented in Baker et al.[18] Many studies have rewritten these calculations into programming languages like R, Python and MATLAB and often distribute them freely on popular code repositories like https://github.com/. An intuitive, open-source and recommendable software package is Quasar (https://quasar.codes/), which is a Python-based data analysis toolbox of the Orange suite[99] extended with an add-on for processing and analysis of FTIR (hyper)spectral data. Custom-written or freely available classification models in Python can easily be imported and applied (e.g., PyChem).[100] For an overview of other freely available chemometric toolboxes, the reader is referred to Morais et al.[84]

4. Conclusions and outlook

Micro-FTIR spectroscopy measurements on culture and sediment-derived single specimen dinocysts from several species were carried out using three major sampling modes (transmission, transflection and ATR) and three different measurement substrates (Au-mirror, Si-window, and ZnSe-window). ATR spectra consistently showed a higher repeatability and reproducibility over spectra collected in nondestructive sampling modes, which typically contain (severe) specimen-specific scattering artifacts and whose quality (mainly SNR) varied with the measurement substrate used. The ATR method proved nearly insensitive to the substrate type and dinocyst geometry and size. The use of strong acids HCl and HF might have a differential impact on the chemistry of the walls from different dinocyst species and should therefore be considered during intercomparison of various samples. Given the diversity of dinoflagellate sample types, specific sample preparation protocols have been proposed to ensure the optimal isolation of individual, clean dinocysts while introducing as little chemical alteration as possible. Depending on the desired research objective, several specific spectral data (pre)processing protocols have been outlined to extract as much qualitative chemical information from acquired micro-FTIR datasets. The proposed and evaluated ATR method can be considered a large methodological step forward in the relatively new field of spectrochemical characterization of dinoflagellates (most notably their cyst walls) and facilitates the collection of robust datasets which more closely reflect the true chemistries of the samples at hand. Ideally in conjunction with other independent and complementary (geo)chemical characterization methods, ATR micro-FTIR spectroscopy could be used to provide answers to major biological and geological standing questions regarding the chemical composition and chemical diversity of dinocysts as well as to evaluate its applicability as a proxy (through geological time).

The use of micro-FTIR spectroscopy in the field of dinoflagellate research is still relatively new. This study is a first attempt in optimizing the methods for collecting, processing, and interpreting FTIR data from single specimen dinocysts by enumeration of relevant variables and based on experimental results. Collection of reliable data

combined with sensible data exploration and interpretation is fundamental for answering standing questions on recent and fossil dinocyst composition and its variability. Following is a selection of currently unresolved questions within the field:

- Can FTIR data from recent dinocysts be used to construct robust chemotaxonomies?
- If so, does this also work for fossil cysts or to what extent does a diagenetic overprint exist?
- Can FTIR spectroscopy be used to gain insights on the transition from bio- to geo-molecules in the walls of dinocyst during burial and diagenesis?
- Can dinocyst chemistry be used as a proxy (e.g., for trophic affinity, temperature, salinity, or UV-flux)?
- To what degree does selective dinocyst preservation exist, what is the link with the depositional environment and what does this imply for the fossil record?

Formulating substantiated answers to these questions will likely require large spectral datasets which can be considered as one of the main challenges as manual picking and measurement of individual specimens is often a tedious and time-consuming task. The methods proposed in this study can also be applied to proof-test, further explore and possibly redefine hypotheses suggested by earlier pioneering micro-FTIR works on dinocysts.[6,8–14,73]

Some FTIR applications on dinoflagellates would require full penetration of the IR light in the sample, which requires measuring in transmission. For example, when characterizing the chemistry of the cell contents of live cells or cysts with the aim of toxin detection. In those cases, it could be helpful to flatten or break open the specimens first to minimize the scattering distortions produced by their irregular morphology. Another alternative would be to extract and homogenize the cell contents and deposit them as very thin layer (~ penetration depth in ATR) on the substrate, which can afterwards be measured in either transmission or ATR.

Apart from chromatography, mass spectrometry-based methods and nuclear magnetic resonance (NMR) other spectroscopy-based methods can be applied in conjunction with micro-FTIR for studying dinocyst chemistry. Since conventional micro-FTIR systems are diffraction-limited to a spatial resolution of about 10–20 μm[101], only whole cyst chemical information can be retrieved. However, by combining visible light with IR light, Optical Photothermal Infrared (O-PTIR) micro-spectroscopy reduces this spatial resolution to about 0.5 μm.[102,103] This allows to measure IR spectra from different locations (e.g., different wall layers, processes, ornamentations) on a single microscopic sample, which can afterwards be merged into hyperspectral chemical maps. O-PTIR can also be used in conjunction with Raman spectroscopy, which utilizes high intensity laser light and is able to reach similar spatial resolutions.[104,105] Raman spectra are typically considered to hold complementary information to FTIR spectra, but no Raman spectroscopy studies have been done on dinocysts so far. Raman spectroscopy only works on samples which exhibit limited autofluorescence. Future research will be focused on exploring and evaluating these methods when applied to dinocysts.

Acknowledgements

Karin Zonneveld is thanked for the coordination of the research stay at MARUM (Bremen University) and the interesting discussions. Coralie Connes, Haifeng Gu, Kazumi Matsuoka, Nathalie Malet, Vera Pospelova, Veronique Séchet and Yoshihito Takano are gratefully acknowledged to provide samples for this study. Eirik Magnussen is thanked for proofreading the manuscript and providing suggestions. We are grateful to the referees for their constructive input.

Declaration of Conflicting Interests

The Author(s) declare(s) that there is no conflict of interest.

Funding

This work was supported by the Hercules Foundation (FWO, Flanders) [FT-IMAGER project – AUGE/13/16].

References

1. M.J. Head. “Modern dinoflagellate cysts and their biological affnities”. In: J. Jansonius, D.C. McGregor, editors. Palynology: principles and applications. American Association of Stratigraphic Palynologists Foundation, 1996. Pp. 1197–1248.
2. N. Lundholm, S. Ribeiro, T.J. Andersen, T. Koch, A. Godhe, F. Ekelund, et al. “Buried alive - Germination of up to a century-old marine protist resting stages”. Phycologia. 2011. 50(6): 629–640. 10.2216/11-16.1.
3. A. Penaud, W. Hardy, C. Lambert, F. Marret, E. Masure, T. Servais, et al. “Dinoflagellate fossils: Geological and biological applications”. Rev. Micropaleontol. Elsevier Masson SAS, 2018. 61(3–4): 235–254. 10.1016/j.revmic.2018.09.003.
4. D.M. Anderson, A.D. Cembella, G.M. Hallegraeff. “Progress in Understanding Harmful Algal Blooms: Paradigm Shifts and New Technologies for Research, Monitoring, and Management”. Ann. Rev. Mar. Sci. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012. 4(1): 143–176. 10.1146/annurev-marine-120308-081121.
5. R.A. Fensome, F.J.R. Taylor, G. Norris, W.A.S. Sarjeant, D.I. Wharton, G.L. Williams. A classification of fossil and living dinoflagellates. Micropaleontology Press Special Paper. no. 7. American Museum of Natural History, New York, 1993.
6. J.P. Kokinos, T.I. Eglinton, M.A. Goñi, J.J. Boon, P.A. Martoglio, D.M. Anderson. “Characterization of a highly resistant biomacromolecular material in the cell wall of a marine dinoflagellate resting cyst”. Org. Geochem. 1998. 28(5): 265–288. 10.1016/S0146-6380(97)00134-4.
7. J.W. De Leeuw, G.J.M. Versteegh, P.F. Van Bergen. “Biomacromolecules of algae and plants and their fossil analogues”. Plant Ecol. 2006. 182(1–2): 209–233. 10.1007/s11258-005-9027-x.
8. G.J.M. Versteegh, P. Blokker, C. Marshall, J. Pross. “Macromolecular composition of the dinoflagellate cyst *Thalassiphora pelagica* (Oligocene, SW Germany)”. Org. Geochem. 2007. 38(10): 1643–1656. 10.1016/j.orggeochem.2007.06.007.
9. G.J.M. Versteegh, P. Blokker, K. Bogus, I.C. Harding, J. Lewis, S. Oltmanns, et al. “Infra red spectroscopy, flash pyrolysis, thermally assisted hydrolysis and

methylation (THM) in the presence of tetramethylammonium hydroxide (TMAH) of cultured and sediment-derived *Lingulodinium polyedrum* (Dinoflagellata) cyst walls". Org. Geochem. Elsevier Ltd, 2012. 43: 92–102. 10.1016/j.orggeochem.2011.10.007.
10. G.J.M. Versteegh, A.J.P. Houben, K.A.F. Zonneveld. "Better molecular preservation of organic matter in an oxic than in a sulfidic depositional environment: Evidence from *Thalassiphora pelagica* (Dinoflagellata, Eocene) cysts". Biogeosciences. 2020. 17(13): 3545–3561. 10.5194/bg-17-3545-2020.
11. K. Bogus, I.C. Harding, A. King, A.J. Charles, K.A.F. Zonneveld, G.J.M. Versteegh. "The composition and diversity of dinosporin in species of the *Apectodinium* complex (Dinoflagellata)". Rev. Palaeobot. Palynol. Elsevier B.V., 2012. 183: 21–31. 10.1016/j.revpalbo.2012.07.001.
12. K. Bogus, K.N. Mertens, J. Lauwaert, I.C. Harding, H. Vrielinck, K.A.F. Zonneveld, et al. "Differences in the chemical composition of organic-walled dinoflagellate resting cysts from phototrophic and heterotrophic dinoflagellates". S. Lin, editor. J. Phycol. 2014. 50(2): 254–266. 10.1111/jpy.12170.
13. P.R. Gurdebeke, K.N. Mertens, V. Pospelova, K. Matsuoka, Z. Li, K.E. Gribble, et al. "Taxonomic revision, phylogeny, and cyst wall composition of the dinoflagellate cyst genus *Votadinium* Reid (Dinophyceae, Peridiniales, Protoperidiniaceae)". Palynology. Taylor & Francis, 2020. 44(2): 310–335. 10.1080/01916122.2019.1580627.
14. P.R. Gurdebeke, K.N. Mertens, P. Meyvisch, K. Bogus, V. Pospelova, S. Louwye. "*Hiddenocysta matsuokae* gen. et sp. nov. from the Holocene of Vancouver Island, British Columbia, Canada". Palynology. Taylor & Francis, 2021. 45(1): 103–114. 10.1080/01916122.2020.1750500.
15. P.R. Gurdebeke. Holocene dinoflagellate cysts and other palynomorphs from Northern Hemisphere estuaries. Ghent University, 2019.
16. P.R. Bunker, P. Jensen. Molecular Symmetry and Spectroscopy. 2nd ed. NRC Research Press, Ottawa, 1998.
17. B.H. Stuart. Infrared Spectroscopy: Fundamentals and Applications. Infrared Spectroscopy: Fundamentals and Applications. 2004. 10.1002/0470011149.
18. M.J. Baker, J. Trevisan, P. Bassan, R. Bhargava, H.J. Butler, K.M. Dorling, et al. "Using Fourier transform IR spectroscopy to analyze biological materials". Nat. Protoc. 2014. 9(8): 1771–1791. 10.1038/nprot.2014.110.
19. P. Lasch, D. Naumann. Infrared Spectroscopy in Microbiology. Encyclopedia of Analytical Chemistry. 2015. 10.1002/9780470027318.a0117.pub2.
20. K. Bogus. The composition and characterization of the organic- walled resting cysts of dinoflagellates : Implications for the preservation of organic matter. Universität Bremen, 2011.
21. A.M. Price, P.R. Gurdebeke, K.N. Mertens, V. Pospelova. "Determining the absolute abundance of dinoflagellate cysts in recent marine sediments III: Identifying the source of *Lycopodium* loss during palynological processing and further testing of the *Lycopodium* marker-grain method". Rev. Palaeobot. Palynol. Elsevier B.V., 2016. 226: 78–90. 10.1016/j.revpalbo.2015.12.009.
22. K.N. Mertens, J. Wolny, C. Carbonell-Moore, K. Bogus, M. Ellegaard, A. Limoges, et al. "Taxonomic re-examination of the toxic armored dinoflagellate *Pyrodinium bahamense* Plate 1906: Can morphology or LSU sequencing separate *P. bahamense* var. *compressum* from var. *bahamense*?" Harmful Algae. 2015. 41: 1–24. 10.1016/j.hal.2014.09.010.
23. K.N. Mertens, Y. Takano, A. Yamaguchi, H. Gu, K. Bogus, A. Kremp, et al. "The molecular characterization of the enigmatic dinoflagellate *Kolkwitziella acuta* reveals an affinity to the Excentrica section of the genus *Protoperidinium*". Syst. Biodivers. 2015. 13(6): 829–844. 10.1080/14772000.2015.1078855.
24. K.N. Mertens, H. Gu, V. Pospelova, N. Chomérat, E. Nézan, P.R. Gurdebeke, et al. "First record of resting cysts of the benthic dinoflagellate *Prorocentrum leve* in a natural reservoir in Gujan-Mestras, Gironde, France". J. Phycol. 2017.

53(6): 1193–1205. 10.1111/jpy.12582.
25. C.J.S. Bolch. “The use of sodium polytungstate for the separation and concentration of living dinoflagellate cysts from marine sediments”. Phycologia. 1997. 36(6): 472–478. 10.2216/i0031-8884-36-6-472.1.
26. E.A. Mousing, T.J. Andersen, M. Ellegaard. “Changes in the Abundance and Species Composition of Phytoplankton in the Last 150 Years in the Southern Black Sea”. Estuaries and Coasts. 2013. 36(6): 1206–1218. 10.1007/s12237-013-9623-2.
27. D. Munsterman, S. Kerstholt. “Sodium polytungstate, a new non-toxic alternative to bromoform in heavy liquid separation”. Rev. Palaeobot. Palynol. 1996. 91(1–4): 417–422. 10.1016/0034-6667(95)00093-3.
28. HYPERION - User Manual. 3rd ed. Bruker Optik GmbH, Ettlingen, 2017.
29. A. Savitzky, M.J.E. Golay. “Smoothing and Differentiation of Data by Simplified Least Squares Procedures”. Anal. Chem. 1964. 36(8): 1627–1639. 10.1021/ac60214a047.
30. B. Zimmermann, A. Kohler. “Optimizing savitzky-golay parameters for improving spectral resolution and quantification in infrared spectroscopy”. Appl. Spectrosc. 2013. 67(8): 892–902. 10.1366/12-06723.
31. N. Colthup, L. Daly, S. Wiberley. Introduction to Infrared and Raman Spectroscopy. Academic Press. 3rd ed. Academic Press, 1990.
32. J. Coates. “Interpretation of Infrared Spectra, A Practical Approach”. Encyclopedia of Analytical Chemistry. John Wiley & Sons, Ltd, Chichester, UK, 2006. Pp. 10815–10837. 10.1002/9780470027318.a5606.
33. P.E. Jardine, F.A.J. Abernethy, B.H. Lomax, W.D. Gosling, W.T. Fraser. “Shedding light on sporopollenin chemistry, with reference to UV reconstructions”. Rev. Palaeobot. Palynol. The Authors, 2017. 238: 1–6. 10.1016/j.revpalbo.2016.11.014.
34. Rs. Team. RStudio; Integrated Development Environment for R. Rstudio, PCB., Boston, MA, 2020.
35. H.J. Gulley-Stahl, S.B. Bledsoe, A.P. Evan, A.J. Sommer. “The Advantages of an Attenuated Total Internal Reflection Infrared Microspectroscopic Imaging Approach for Kidney Biopsy Analysis”. Appl. Spectrosc. 2010. 64(1): 15–22. 10.1366/000370210792966161.
36. J. Lee, E. Gazi, J. Dwyer, M. Brown, N. Clarke, J.M. Nicholson, et al. “Optical artefacts in transflection mode FTIR microspectroscopic images of single cells on a biological support: The effect of back-scattering into collection optics”. Analyst. 2007. 132(8): 750–755. 10.1039/b702064c.
37. A. Kohler, J. Sulé-Suso, G.D. Sockalingum, M. Tobin, F. Bahrami, Y. Yang, et al. “Estimating and Correcting Mie Scattering in Synchrotron-Based Microscopic Fourier Transform Infrared Spectra by Extended Multiplicative Signal Correction”. Appl. Spectrosc. 2008. 62(3): 259–266. 10.1366/000370208783759669.
38. P. Bassan, H.J. Byrne, F. Bonnier, J. Lee, P. Dumas, P. Gardner. “Resonant Mie scattering in infrared spectroscopy of biological materials - Understanding the ‘dispersion artefact’”. Analyst. 2009. 134(8): 1586–1593. 10.1039/b904808a.
39. H. Brooke, B. V. Bronk, J.N. McCutcheon, S.L. Morgan, M.L. Myrick. “A study of electric field standing waves on reflection microspectroscopy of polystyrene particles.” Appl. Spectrosc. 2009. 63(11): 1293–1302. 10.1366/000370209789806902.
40. P. Bassan. Light scattering during infrared spectroscopic measurements of biomedical samples. University of Manchester, 2011.
41. P. Bassan, J. Lee, A. Sachdeva, J. Pissardini, K.M. Dorling, J.S. Fletcher, et al. “The inherent problem of transflection-mode infrared spectroscopic microscopy and the ramifications for biomedical single point and imaging applications”. Analyst. 2013. 138(1): 144–157. 10.1039/c2an36090j.
42. T. Konevskikh, A. Ponossov, R. Blümel, R. Lukacs, A. Kohler. “Fringes in FTIR spectroscopy revisited: Understanding and modelling fringes in infrared

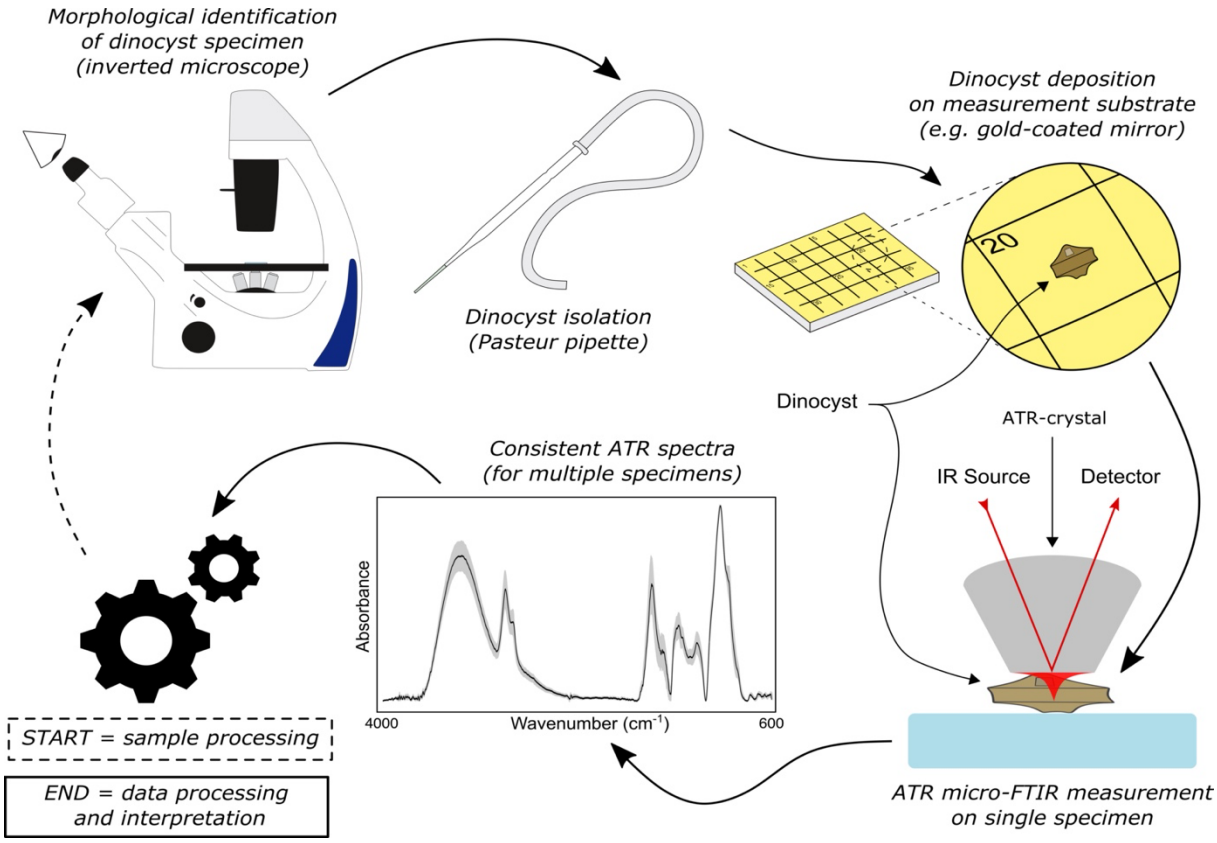
spectroscopy of thin films". Analyst. Royal Society of Chemistry, 2015. 140(12): 3969–3980. 10.1039/c4an02343a.
43. M.J. Pilling, P. Bassan, P. Gardner. "Comparison of transmission and transflectance mode FTIR imaging of biological tissue". Analyst. Royal Society of Chemistry, 2015. 140(7): 2383–2392. 10.1039/c4an01975j.
44. E. Staniszewska-Slezak, A. Rygula, K. Malek, M. Baranska. "Transmission versus transflection mode in FTIR analysis of blood plasma: Is the electric field standing wave effect the only reason for observed spectral distortions?" Analyst. Royal Society of Chemistry, 2015. 140(7): 2412–2421. 10.1039/c4an01842g.
45. R. Blümel, M. Bağcioğlu, R. Lukacs, A. Kohler. "Infrared refractive index dispersion of polymethyl methacrylate spheres from Mie ripples in Fourier-transform infrared microscopy extinction spectra". J. Opt. Soc. Am. A. 2016. 33(9): 1687. 10.1364/josaa.33.001687.
46. R. Blümel, R. Lukacs, B. Zimmermann, M. Bağcıoğlu, A. Kohler. "Observation of Mie ripples in the synchrotron Fourier transform infrared spectra of spheroidal pollen grains". J. Opt. Soc. Am. A. 2018. 35(10): 1769. 10.1364/josaa.35.001769.
47. T.G. Mayerhöfer, J. Popp. "Electric field standing wave effects in internal reflection and ATR spectroscopy". Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. Elsevier B.V., 2018. 191: 165–171. 10.1016/j.saa.2017.10.007.
48. T.G. Mayerhöfer, J. Popp. "The electric field standing wave effect in infrared transflection spectroscopy". Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 2018. 191: 283–289. 10.1016/j.saa.2017.10.033.
49. M.A. Brandsrud, R. Blümel, J.H. Solheim, E.A. Magnussen, E. Seim, A. Kohler. Does chaotic scattering affect the extinction efficiency in quasi-spherical scatterers? In: J. Popp, C. Gergely, editors. Biomedical Spectroscopy, Microscopy, and Imaging. SPIE, 2020. P. 11. 10.1117/12.2556390.
50. M.A. Brandsrud, R. Blümel, J.H. Solheim, A. Kohler. "The effect of deformation of absorbing scatterers on Mie-type signatures in infrared microspectroscopy". Sci. Rep. Nature Publishing Group UK, 2021. 11(1): 1–14. 10.1038/s41598-021-84064-5.
51. H. Martens, E. Stark. "Extended multiplicative signal correction and spectral interference subtraction: New preprocessing methods for near infrared spectroscopy". J. Pharm. Biomed. Anal. 1991. 9(8): 625–635. 10.1016/0731-7085(91)80188-F.
52. M. Romeo, M. Diem. "Correction of dispersive line shape artifact observed in diffuse reflection infrared spectroscopy and absorption/reflection (transflection) infrared micro-spectroscopy". Vib. Spectrosc. 2005. 38(1–2): 129–132. 10.1016/j.vibspec.2005.04.003.
53. N.K. Afseth, A. Kohler. "Extended multiplicative signal correction in vibrational spectroscopy, a tutorial". Chemom. Intell. Lab. Syst. Elsevier B.V., 2012. 117: 92–99. 10.1016/j.chemolab.2012.03.004.
54. P. Bassan, A. Kohler, H. Martens, J. Lee, H.J. Byrne, P. Dumas, et al. "Resonant Mie Scattering (RMieS) correction of infrared spectra from highly scattering biological samples". Analyst. 2010. 135(2): 268–277. 10.1039/b921056c.
55. P. Bassan, A. Kohler, H. Martens, J. Lee, E. Jackson, N. Lockyer, et al. "RMieS-EMSC correction for infrared spectra of biological cells: Extension using full Mie theory and GPU computing". J. Biophotonics. 2010. 3(8–9): 609–620. 10.1002/jbio.201000036.
56. T. Konevskikh, R. Lukacs, A. Kohler. "An improved algorithm for fast resonant Mie scatter correction of infrared spectra of cells and tissues". J. Biophotonics. 2018. 11(1): 1–10. 10.1002/jbio.201600307.
57. J. Skogholt, K.H. Liland, U.G. Indahl. "Preprocessing of spectral data in the extended multiplicative signal correction framework using multiple reference spectra". J. Raman Spectrosc. 2019. 50(3): 407–417. 10.1002/jrs.5520.
58. A.J. Schofield, R. Blümel, A. Kohler, R. Lukacs, C.J. Hirschmugl. "Extracting

pure absorbance spectra in infrared microspectroscopy by modeling absorption bands as Fano resonances". J. Chem. Phys. 2019. 150(15). 10.1063/1.5085207.
59. J.H. Solheim, E. Gunko, D. Petersen, F. Großerueschkamp, K. Gerwert, A. Kohler. "An open-source code for Mie extinction extended multiplicative signal correction for infrared microscopy spectra of cells and tissues". J. Biophotonics. 2019. 12(8): 1–14. 10.1002/jbio.201800415.
60. J. Yang, J. Xu, X. Zhang, C. Wu, T. Lin, Y. Ying. "Deep learning for vibrational spectral analysis: Recent progress and a practical guide". Anal. Chim. Acta. Elsevier Ltd, 2019. 1081: 6–17. 10.1016/j.aca.2019.06.012.
61. S. Guo, T.G. Mayerhöfer, S. Pahlow, U. Hübner, J. Popp, T. Bocklitz. "Deep learning for 'artefact' removal in infrared spectroscopy". Analyst. 2020. 145(15): 5213–5220. 10.1039/d0an00917b.
62. P. Pradhan, S. Guo, O. Ryabchykov, J. Popp, T.W. Bocklitz. "Deep learning a boon for biophotonics?" J. Biophotonics. 2020. 13(6): 1–24. 10.1002/jbio.201960186.
63. A.P. Raulf, J. Butke, L. Menzen, C. Küpper, F. Großerueschkamp, K. Gerwert, et al. "A representation learning approach for recovering scatter-corrected spectra from Fourier-transform infrared spectra of tissue samples". J. Biophotonics. 2021. 14(3): 1–8. 10.1002/jbio.202000385.
64. A.P. Raulf, J. Butke, L. Menzen, C. Küpper, F. Großerueschkamp, K. Gerwert, et al. "Deep neural networks for the correction of mie scattering in fourier-transformed infrared spectra of biological samples". arXiv. 2020.
65. A. Gupper, P. Wilhelm, M. Schmied, S.G. Kazarian, K.L.A. Chan, J. Reußner. "Combined application of imaging methods for the characterization of a polymer blend". Appl. Spectrosc. 2002. 56(12): 1515–1523. 10.1366/000370202321115959.
66. B. Zimmermann, Z. Tkalčec, A. Mešić, A. Kohler. "Characterizing aeroallergens by infrared spectroscopy of fungal spores and pollen". PLoS One. 2015. 10(4): 1–22. 10.1371/journal.pone.0124240.
67. K.L. Andrew Chan, S.G. Kazarian. "Attenuated total reflection Fourier-transform infrared (ATR-FTIR) imaging of tissues and live cells". Chem. Soc. Rev. Royal Society of Chemistry, 2016. 45(7): 1850–1864. 10.1039/c5cs00515a.
68. K.L.A. Chan, S.G. Kazarian. "New opportunities in micro- and macro-attenuated total reflection infrared spectroscopic imaging: Spatial resolution and sampling versatility". Appl. Spectrosc. 2003. 57(4): 381–389. 10.1366/00037020360625907.
69. Z. Li, P.M. Fredericks, L. Rintoul, C.R. Ward. "Application of attenuated total reflectance micro-Fourier transform infrared (ATR-FTIR) spectroscopy to the study of coal macerals: Examples from the Bowen Basin, Australia". Int. J. Coal Geol. 2007. 70(1-3 SPEC. ISS.): 87–94. 10.1016/j.coal.2006.01.006.
70. G. Kosa, V. Shapaval, A. Kohler, B. Zimmermann. "FTIR spectroscopy as a unified method for simultaneous analysis of intra- and extracellular metabolites in high-throughput screening of microbial bioprocesses". Microb. Cell Fact. BioMed Central, 2017. 16(1): 1–11. 10.1186/s12934-017-0817-3.
71. V. Shapaval, J. Brandenburg, J. Blomqvist, V. Tafintseva, V. Passoth, M. Sandgren, et al. "Biochemical profiling, prediction of total lipid content and fatty acid profile in oleaginous yeasts by FTIR spectroscopy". Biotechnol. Biofuels. BioMed Central, 2019. 12(1): 1–12. 10.1186/s13068-019-1481-0.
72. S. Diehn, B. Zimmermann, V. Tafintseva, M. Bağcıoğlu, A. Kohler, M. Ohlson, et al. "Discrimination of grass pollen of different species by FTIR spectroscopy of individual pollen grains". Anal. Bioanal. Chem. 2020. 10.1007/s00216-020-02628-2.
73. P.R. Gurdebeke, K.N. Mertens, K. Bogus, F. Marret, N. Chomérat, H. Vrielinck, et al. "Taxonomic Re-Investigation and Geochemical Characterization of Reid's (1974) Species of *Spiniferites* from Holotype and Topotype Material". Palynology. 2018. 42(sup1): 93–110. 10.1080/01916122.2018.1465735.

74. C. Rumpel, N. Rabia, S. Derenne, K. Quenea, K. Eusterhues, I. Kögel-Knabner, et al. “Alteration of soil organic matter following treatment with hydrofluoric acid (HF)”. Org. Geochem. 2006. 37(11): 1437–1451. 10.1016/j.orggeochem.2006.07.001.
75. Y. Audette, D.S. Smith, C.T. Parsons, J. Longstaffe, W. Chen, F. Rezanezhad, et al. “Impact of Hydrofluoric Acid Treatment on Humic Acid Properties Extracted from Organic Soils and an Organic Amendment: A Technical Evaluation”. Soil Sci. Soc. Am. J. 2019. 83(4): 1219–1226. 10.2136/sssaj2018.11.0419.
76. D.M. Anderson, J.J. Lively, E.M. Reardon, C.A. Price. “Sinking characteristics of dinoflagellate cysts”. Limnol. Oceanogr. 1985. 30(5): 1000–1009. 10.4319/lo.1985.30.5.1000.
77. D.M. Considine. Van Nostrand’s Scientific Encyclopedia. 7th ed. Van Nostrand Reinhold, New York, 1989.
78. S.W. Bruun, A. Kohler, I. Adt, G.D. Sockalingum, M. Manfait, H. Martens. “Correcting attenuated total reflection-fourier transform infrared spectra for water vapor and carbon dioxide”. Appl. Spectrosc. 2006. 60(9): 1029–1039. 10.1366/000370206778397371.
79. P. Lasch. “Spectral pre-processing for biomedical vibrational spectroscopy and microspectroscopic imaging”. Chemom. Intell. Lab. Syst. 2012. 117(May): 100–114. 10.1016/j.chemolab.2012.03.011.
80. V. Tafintseva, V. Shapaval, M. Smirnova, A. Kohler. “Extended multiplicative signal correction for FTIR spectral quality test and pre-processing of infrared imaging data”. J. Biophotonics. 2020. 13(3): 1–15. 10.1002/jbio.201960112.
81. F. Holler, D.H. Burns, J.B. Callis. “Direct use of second derivatives in curve-fitting procedures”. Appl. Spectrosc. 1989. 43(5): 877–882. 10.1366/0003702894202292.
82. J. Trevisan, P.P. Angelov, P.L. Carmichael, A.D. Scott, F.L. Martin. “Extracting biological information with computational analysis of Fourier-transform infrared (FTIR) biospectroscopy datasets: Current practices to future perspectives”. Analyst. 2012. 137(14): 3202–3215. 10.1039/c2an16300d.
83. H.J. Butler, B.R. Smith, R. Fritzsch, P. Radhakrishnan, D.S. Palmer, M.J. Baker. “Optimised spectral pre-processing for discrimination of biofluids via ATR-FTIR spectroscopy”. Analyst. Royal Society of Chemistry, 2018. 143(24): 6121–6134. 10.1039/c8an01384e.
84. C.L.M. Morais, K.M.G. Lima, M. Singh, F.L. Martin. “Tutorial: multivariate classification for vibrational spectroscopy in biological samples”. Nat. Protoc. Springer US, 2020. 15(7): 2143–2162. 10.1038/s41596-020-0322-8.
85. M.A. Al-Holy, M. Lin, H. Al-Qadiri, A.G. Cavinato, B.A. Rasco. “Classification of foodborne pathogens by Fourier transform infrared spectroscopy and pattern recognition techniques”. J. Rapid Methods Autom. Microbiol. 2006. 14(2): 189–200. 10.1111/j.1745-4581.2006.00045.x.
86. R. Dell’Anna, P. Lazzeri, M. Frisanco, F. Monti, F. Malvezzi Campeggi, E. Gottardini, et al. “Pollen discrimination and classification by Fourier transform infrared (FT-IR) microspectroscopy and machine learning”. Anal. Bioanal. Chem. 2009. 394(5): 1443–1452. 10.1007/s00216-009-2794-9.
87. J. Ollesch, S.L. Drees, H.M. Heise, T. Behrens, T. Brüning, K. Gerwert. “FTIR spectroscopy of biofluids revisited: An automated approach to spectral biomarker identification”. Analyst. 2013. 138(14): 4092–4102. 10.1039/c3an00337j.
88. A.C.M. Julier, P.E. Jardine, A.L. Coe, W.D. Gosling, B.H. Lomax, W.T. Fraser. “Chemotaxonomy as a tool for interpreting the cryptic diversity of Poaceae pollen”. Rev. Palaeobot. Palynol. The Authors, 2016. 235: 140–147. 10.1016/j.revpalbo.2016.08.004.
89. M. Bağcıoğlu, A. Kohler, S. Seifert, J. Kneipp, B. Zimmermann. “Monitoring of plant–environment interactions by high-throughput FTIR spectroscopy of pollen”. Methods Ecol. Evol. 2017. 8(7): 870–880. 10.1111/2041-210X.12697.
90. A. Woutersen, P.E. Jardine, R.G. Bogotá-Angel, H.X. Zhang, D. Silvestro, A.

Antonelli, et al. “A novel approach to study the morphology and chemistry of pollen in a phylogenetic context, applied to the halophytic taxon Nitraria L.(Nitrariaceae)”. PeerJ. 2018. 2018(7): 1–31. 10.7717/peerj.5055.
91. B. Zimmermann. “Chemical characterization and identification of Pinaceae pollen by infrared microspectroscopy”. Planta. Springer Berlin Heidelberg, 2018. 247(1): 171–180. 10.1007/s00425-017-2774-9.
92. F. Muthreich, B. Zimmermann, H. John B Birks, C.M. Vila-Viçosa, A.W.R. Seddon. “Chemical variations in *Quercus* pollen as a tool for taxonomic identification: Implications for long-term ecological and biogeographical research”. bioRxiv. 2019. (September). 10.1101/761148.
93. E. Korb, M. Bağcıoğlu, E. Garner-Spitzer, U. Wiedermann, M. Ehling-Schulz, I. Schabussova. “Machine learning-empowered ftir spectroscopy serum analysis stratifies healthy, allergic, and sit-treated mice and humans”. Biomolecules. 2020. 10(7): 1–17. 10.3390/biom10071058.
94. J. Grdadolnik. “ATR-FTIR spectroscopy: it’s advantages and limitations”. Acta Chim. Slov. 2002. 49(3): 631–642.
95. T.G. Mayerhöfer, S. Pahlow, U. Hübner, J. Popp. “CaF2: An Ideal Substrate Material for Infrared Spectroscopy?” Anal. Chem. 2020. 92(13): 9024–9031. 10.1021/acs.analchem.0c01158.
96. H. Martens, J.P. Nielsen, S.B. Engelsen. “Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures”. Anal. Chem. 2003. 75(3): 394–404. 10.1021/ac020194w.
97. I.L. Rasskazov, R. Singh, P.S. Carney, R. Bhargava. “Extended Multiplicative Signal Correction for Infrared Microspectroscopy of Heterogeneous Samples with Cylindrical Domains”. Appl. Spectrosc. 2019. 73(8): 859–869. 10.1177/0003702819844528.
98. T.G. Mayerhöfer, S. Pahlow, J. Popp. “The Bouguer-Beer-Lambert Law: Shining Light on the Obscure”. ChemPhysChem. 2020. 10.1002/cphc.202000464.
99. J. Demsar, T. Curk, A. Erjavec, C. Gorup, T. Hocevar, M. Milutinovic, et al. “Orange: Data Mining Toolbox in Python”. J. Mach. Learn. Res. 2013. 14: 2349–2353.
100. R.M. Jarvis, D. Broadhurst, H. Johnson, N.M. O’Boyle, R. Goodacre. “PYCHEM: A multivariate analysis package for python”. Bioinformatics. 2006. 22(20): 2565–2566. 10.1093/bioinformatics/btl416.
101. M. Milosevic. Internal Reflection and ATR Spectroscopy. John Wiley & Sons, Inc., Hoboken, NJ, USA, 2012. 10.1002/9781118309742.
102. J.A. Reffner. “Advances in Infrared Microspectroscopy and Mapping Molecular Chemical Composition at Submicrometer Spatial Resolution”. Spectroscopy. 2018. 33(9): 12–17.
103. N. Baden, H. Kobayashi, N. Urayama. “Submicron-resolution polymer orientation mapping by optical photothermal infrared spectroscopy”. Int. J. Polym. Anal. Charact. Taylor & Francis, 2020. 25(1): 1–7. 10.1080/1023666X.2020.1735851.
104. N.E. Olson, Y. Xiao, Z. Lei, A.P. Ault. “Simultaneous Optical Photothermal Infrared (O-PTIR) and Raman Spectroscopy of Submicrometer Atmospheric Particles”. Anal. Chem. 2020. 92(14): 9932–9939. 10.1021/acs.analchem.0c01495.
105. A. Spadea, J. Denbigh, M.J. Lawrence, M. Kansiz, P. Gardner. “Analysis of Fixed and Live Single Cells Using Optical Photothermal Infrared with Concomitant Raman Spectroscopy”. Anal. Chem. 2021. 10.1021/acs.analchem.0c04846.

Figures



Graphical abstract

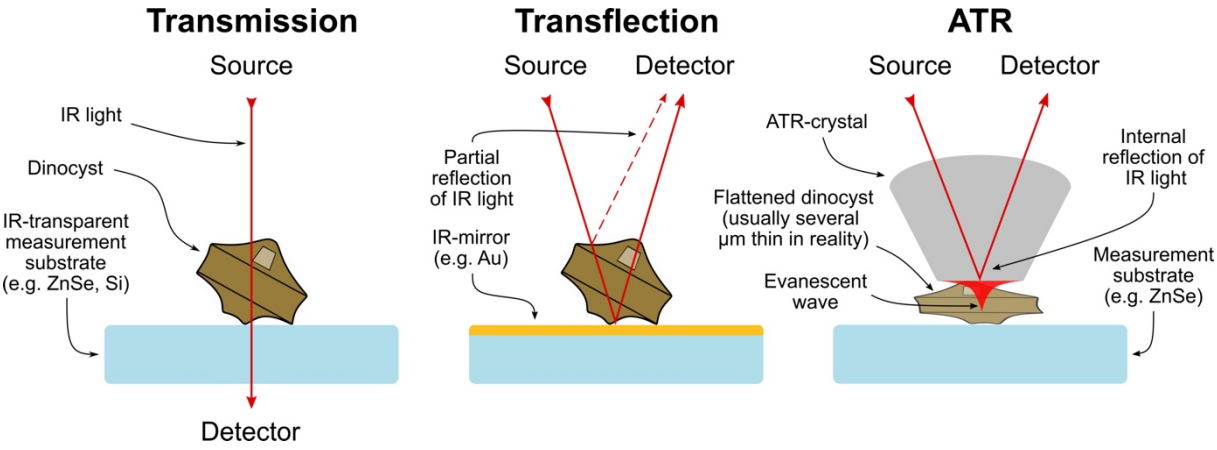


Figure 1. Schematic representation of the three main sampling modes for FTIR spectroscopy applied to single specimen of *Selenopemphix nephroides*, microscopic dinocysts.

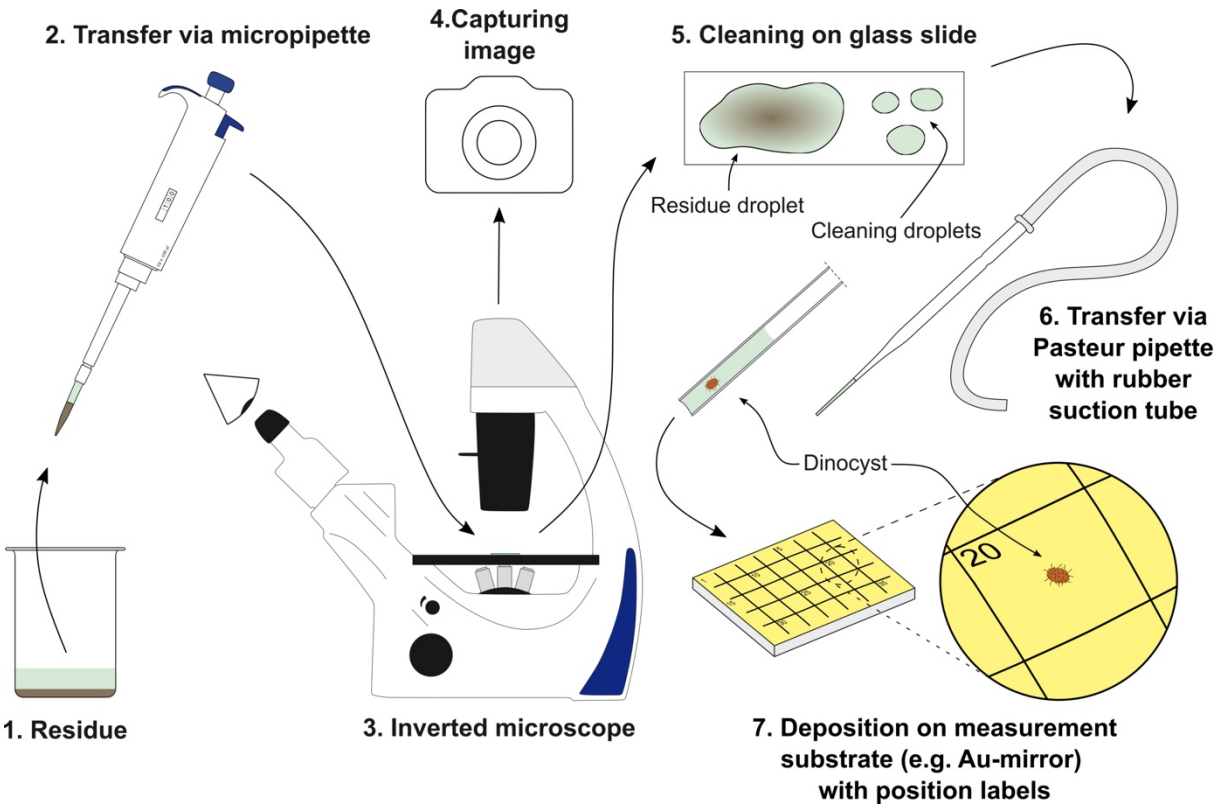


Figure 2. Graphic representation of the isolation, cleaning, and transfer procedure of individual dinocysts prior to micro-FTIR analysis. Adapted from Gurdebeke[15] with permission from the author.

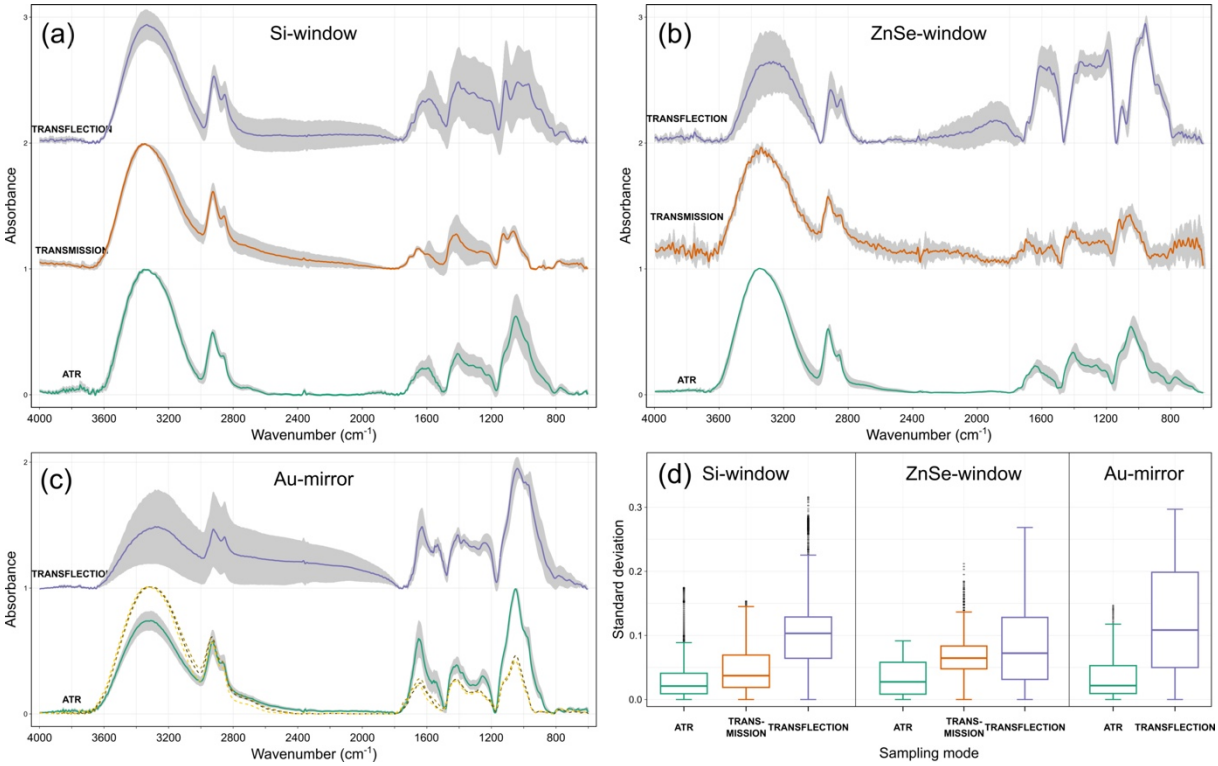


Figure 3. Micro-FTIR measurements of *Lingulodinium machaerophorum* cysts using multiple sampling modes and measurement substrates: (a) Si-window, (b) ZnSe-window and (c) Au-coated mirror. The full lines represent the mean culture-derived *L. machaerophorum* spectrum from transmission, transflection and ATR data, as measured on the respective substrates. The grey area surrounding the mean spectra visualizes the mean standard deviation at each recorded wavenumber. (c) The dashed lines represent sediment-derived *L. machaerophorum* spectra. (d) Box plots showing the mean standard deviations from the mean culture-derived *L. machaerophorum* spectra for each of the measurement substrates and sampling modes. The box encompasses Q1–Q3, the whiskers represent the minimum and maximum values and outliers are represented by individual dots.

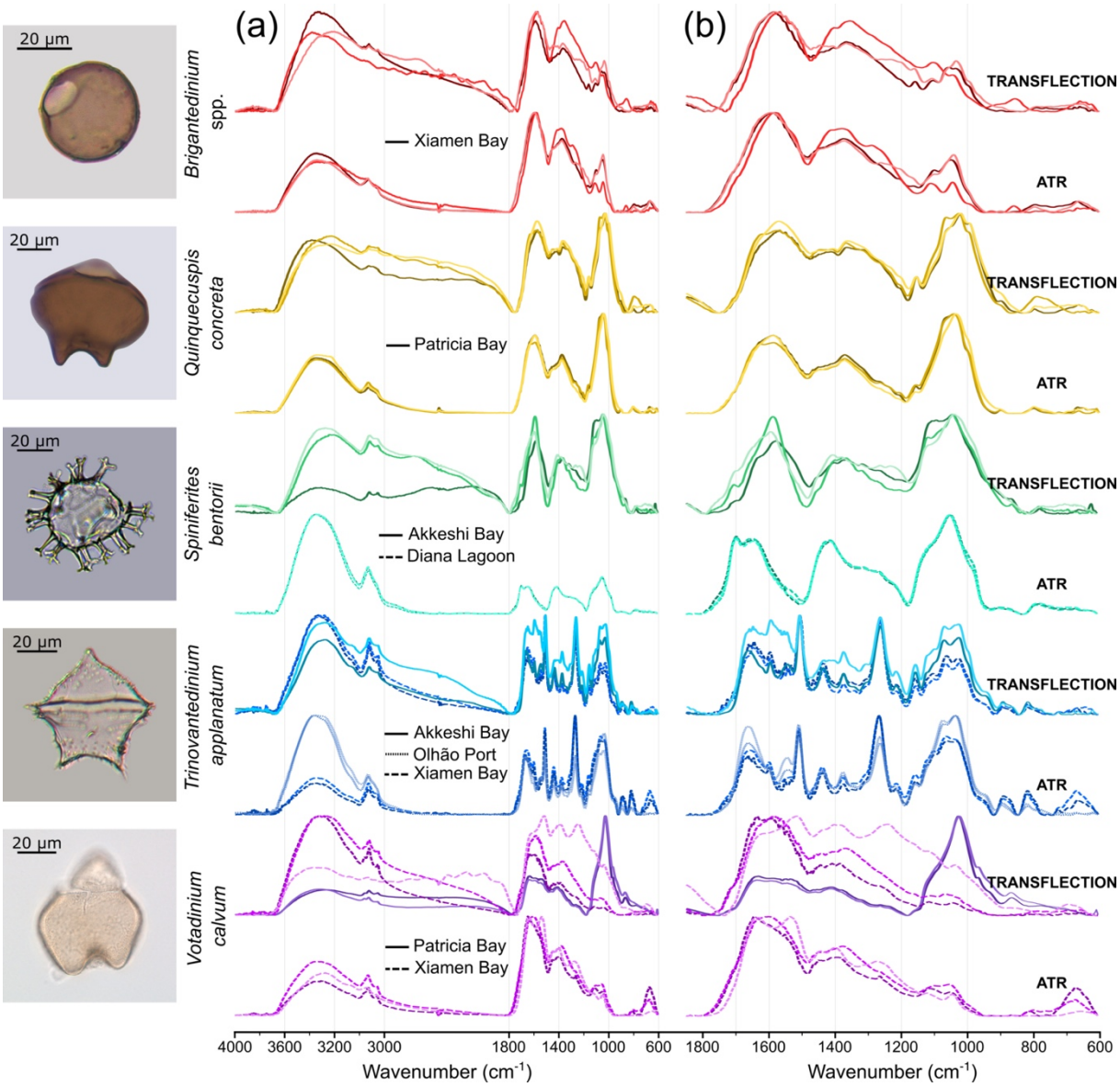


Figure 4. Comparison of micro-FTIR transflection and ATR measurements on an Au-mirror measurement substrate, for morphologically different dinocyst species (photographs and scale bar on the left). (a) full spectra and (b) the fingerprint region, both min-max normalized over their respective extent. Line types indicate sample locations.

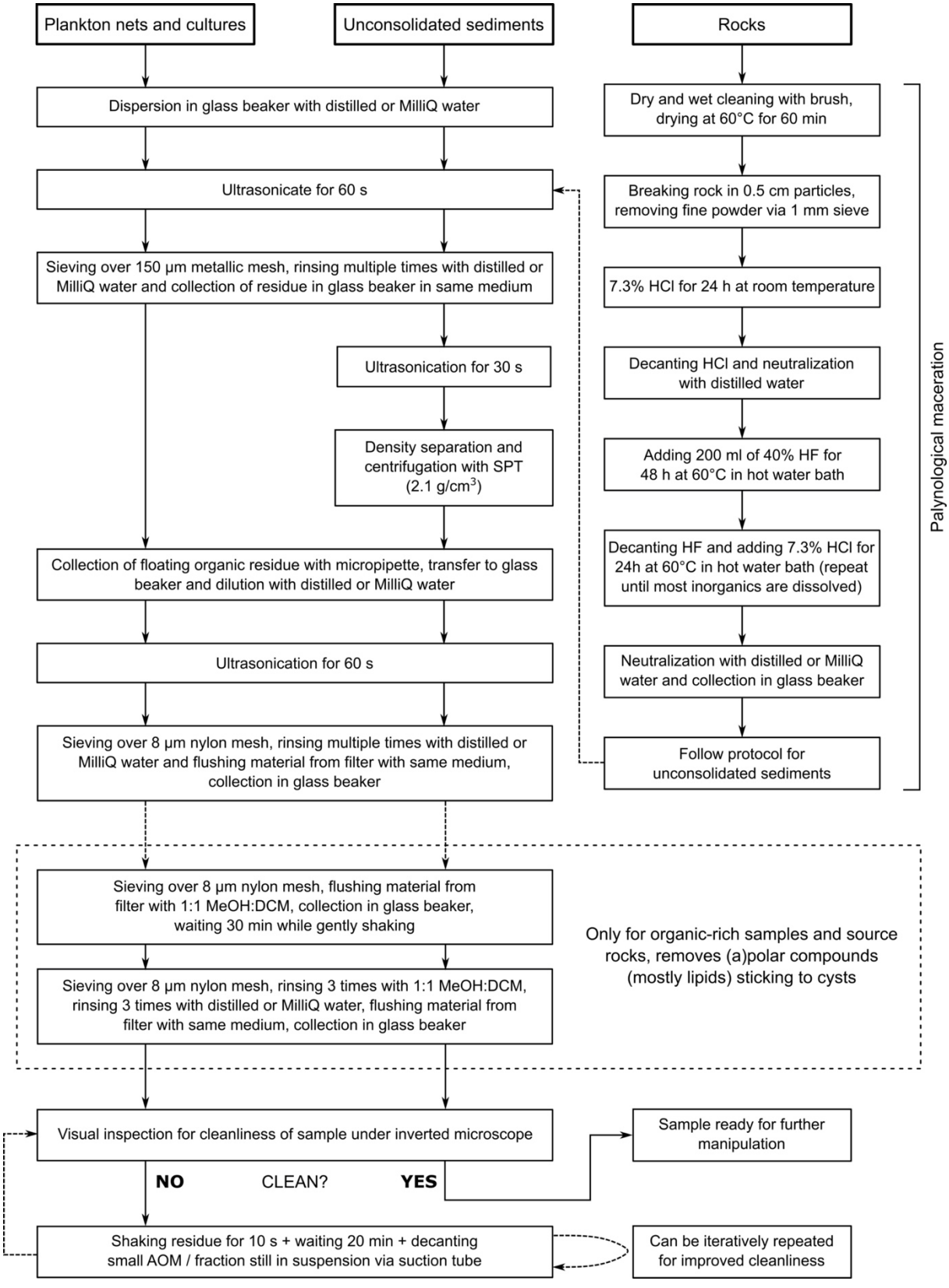


Figure 5. Flowchart showing proposed dinocyst sample preparation protocols, depending on the sample archetype at hand. AOM = amorphous organic matter.

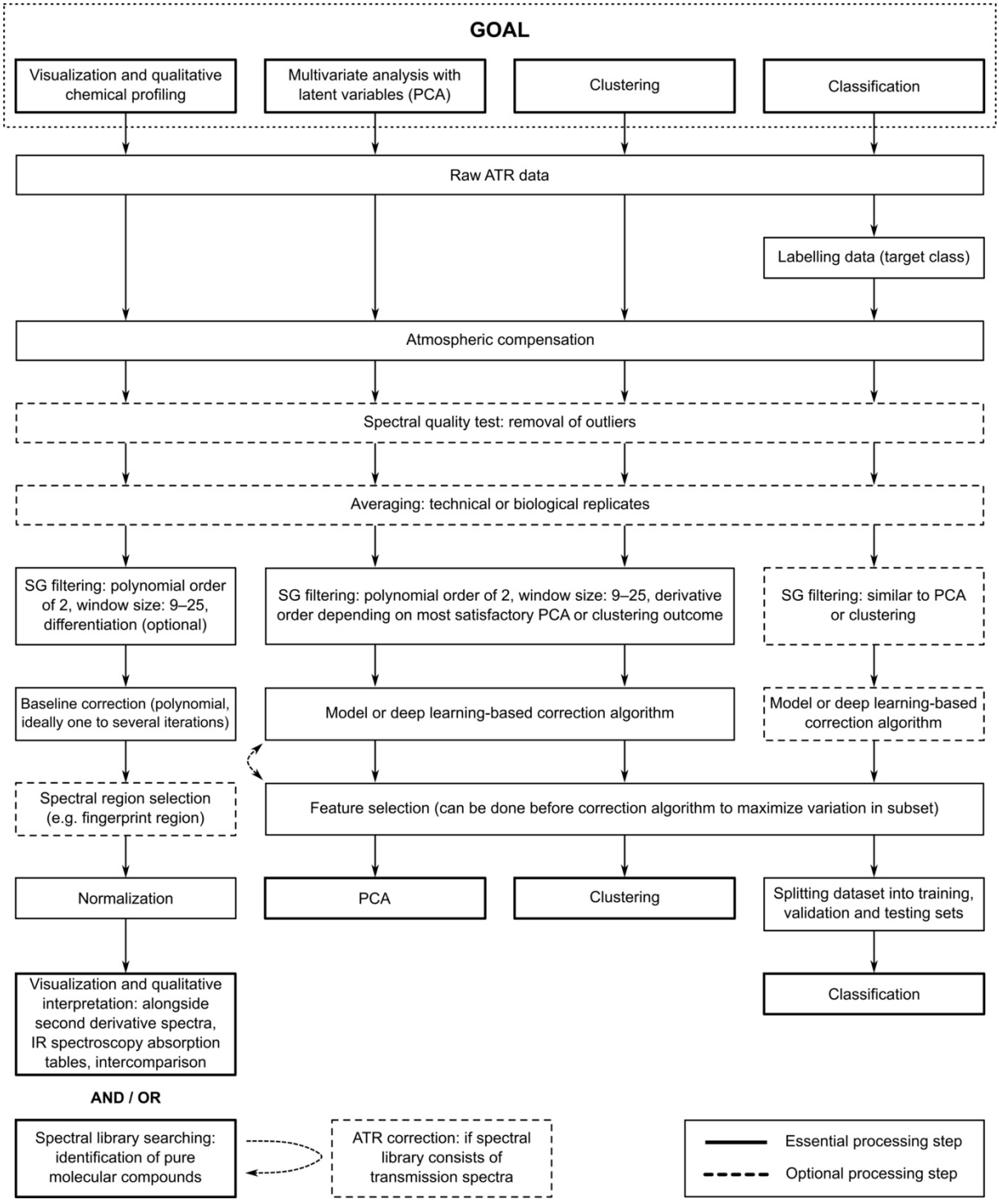


Figure 6. Flowchart showing proposed single specimen dinocyst ATR micro-FTIR data processing steps, depending on the desired analysis and research goals. PCA = principal component analysis.

Tables

Location	Akkeshi Bay (Hokkaido, Japan)	Uvic 110-3XX Diana Lagoon (Corsica, France)	Uvic 110-3XX Diana Lagoon (Corsica, France)	Station 1, Olhão Port (Algarve, Portugal)	Station B, Patricia Bay, Saanich Inlet (Vancouver Island, Canada)	Xiamen Bay (China)
Latitude	43°01.013' N	42°07.66' N	42°07.66' N	37°01.3572' N	48°39.150' N	24°35.568' N
Longitude	144°51.826' E	9°31.72' E	9°31.72' E	7°50.2303' W	123°26.976' W	118°9.198' E
Water depth (m)	1	1	1	3	9.8	10
Sampling method	TFO corer	Hand sampling	Hand sampling	Petite Ponar grab	Van Veen grab	Van Veen grab
Sampling date	20/07/2011	10/01/2018	18/01/2016	8/10/2019	19/05/2009	27/02/2018
Sample type	Sediment	Culture (at 21 °C)	Sediment	Sediment	Sediment	Sediment
Processing steps (in downwards order of occurrence)	Palynological maceration	Disperse 5 ml of culture (Milli-Q)	Disperse several g of sediment (Milli-Q)	Disperse several g of sediment (distilled water)	Disperse several g of sediment (distilled water)	Palynological maceration
	Filter (125+ 8 µm)	Ultrasonicate (60s)	Ultrasonicate (60s)	Ultrasonicate (30s)	Ultrasonicate (60s)	Filter (125+ 8 µm)
	Disperse several g of residue (Miili-Q)	Filter (8 µm)	Filter (125+ 8 µm)	Filter (125+ 8 µm)	Filter (125+ 8 µm)	Disperse several g of residue (Miili-Q)
	Ultrasonicate (30s)	Collect residue in glass vial (Milli-Q)	Ultrasonicate (30s)	Rinse with organic solvents	Collect residue in glass vial (distilled water)	Ultrasonicate (30s)
	Filter (20 µm)		Density separation: SPT (2.1 g/ml)	Ultrasonicate (60s)	Fixate (EtOH)	Filter (20 µm)
	Rinse with organic solvents		Ultrasonicate (60s)	Density separation: SPT (2.1 g/ml)		Rinse with organic solvents
	Ultrasonicate (30s)		Filter (8 µm)	Filter (5 µm)		Ultrasonicate (30s)
	Filter (20 µm)		Collect residue in glass vial (Milli-Q)	Collect residue in glass vial (Milli-Q)		Filter (20 µm)
	Collect residue in glass vial (Milli-Q)		Fixate (EtOH)	Fixate (EtOH)		Collect residue in glass vial (Milli-Q)

Table 1. Sediment and culture samples from which dinocysts were isolated and analyzed: background information and sample processing steps.

Taxon	Location	Sample type	Measurement substrate	Palynological treatment	micro-FTIR sampling mode + no. of spectra		
					Transmission	Transflection	ATR
Brigantedinium spp.	Xiamen Bay	sediment	Au-mirror	**yes**		3	3
Lingulodinium machaerophorum	Diana Lagoon	culture	Au-mirror	no		10	10
Lingulodinium machaerophorum	Diana Lagoon	sediment	Au-mirror	no			3
Lingulodinium machaerophorum	Diana Lagoon	culture	ZnSe-window	no	10	10	10
Lingulodinium machaerophorum	Diana Lagoon	culture	Si-window	no	5	5	5
Quinquecuspis concreta	Patricia Bay	sediment	Au-mirror	no		3	3
Spiniferites bentorii	Diana Lagoon	sediment	Au-mirror	no			3
Spiniferites bentorii	Akkeshi Bay	sediment	Au-mirror	**yes**		3	
Trinovantedinium applanatum	Xiamen Bay	sediment	Au-mirror	**yes**		2	2
Trinovantedinium applanatum	Akkeshi Bay	sediment	Au-mirror	**yes**		2	
Trinovantedinium applanatum	Olhão Port	sediment	Au-mirror	no			2
Votadinium calvum	Xiamen Bay	sediment	Au-mirror	**yes**		3	3
Votadinium calvum	Patricia Bay	sediment	Au-mirror	no		2	

Table 2. Dinocyst species and number of specimens per micro-FTIR sampling mode and measurement substrate analyzed in this study.

Region	Group frequency, Wavenumber (cm^{-1})	Origin	Assignment
High wavenumber region	3550–3200 (broad)	O-H	OH stretching
	3000–2800 (medium, several)	C-H	Aliphatic CH and CH_2 (a)symmetric stretching
Fingerprint region A.	1750-1700 (weak to medium)	C=O	Lipid ester and carboxyl stretching
	1700-1600 (medium to strong)	C=O	Amide I (protein) stretching
	1680–1620 (medium)	C=C	Alkenyl C=C stretch
	1650-1550 (medium to strong)	N-H, C-N	Amide II (protein) N-H bending, C-N stretching
	1600 (medium)	C=C	Conjugated C=C stretch
	1555 (weak to medium)	N-O	NO stretch
	1510 (medium)	C=C	Aromatic C=C stretching
Fingerprint region B.	1450–1350 (medium, several)	C-H	Aliphatic CH and CH_2 bending
	1350-1200 (weak to medium)	N-H, C-N	Amide III (protein) N-H bending, C-N stretching
	1335 (weak)	O-H	OH in-plane bending
	1315 (weak)	C-H	CH_2 wagging
	1203 (weak)	C-O	Asymmetric ester stretching
Fingerprint region C.	1160 (medium)	C-O-C	Asymmetric ether stretching
	1110 (medium)	C-O	Glucose ring stretch
	1060 (strong)	C-OR	Anhydroglucose ring stretching
	1030 (strong)	C-O-C	Pyranose ring skeletal vibration
	1000-985 (medium, several)	C=C	Monosubstituted alkene bending
	897 (weak)	C-H	CH wag of β-glycosidic bond
	860 (weak)	C-H	Aromatic C-H deformation
	715 + 615 (weak)	C-H	$(CH_2)_n$-rocking ($n \geq 3$)
	665 (weak)	C-OH	Out-of-plane bending (Fan et al. 2012)

Table 3. Common absorption bands in micro-FTIR spectra from individual dinocysts. The fingerprint region (1750-600 cm^{-1}) is divided in three subregions, based on the presence of characteristic groups of absorptions.

Mode	Pathway of IR light	Measurement substrate	Measurement area (µm)	Advantages	Limitations
Transmission	Travels one time through sample and IR-transparent substrate and is detected at opposite side of the source.	Potassium bromide, calcium or barium fluoride, zinc selenide, silicon.	20 × 20 to 100 × 100 (adaptable, square)	Nondestructive Full sample penetration Adaptable aperture Less complex scattering effects than for transflection	Scattering artifacts (correctable to a certain degree) Lower SNR than ATR Minimal sample thickness for sufficient absorption Maximal sample thickness limit for sufficient signal collection
Transflection	Travels twice through sample as it is reflected on an IR-reflective substrate and is detected at the same side as the source.	Calcium or barium fluoride, zinc selenide, MirrIR low-E coated glass	20 × 20 to 100 × 100 (adaptable, square)	Nondestructive Full sample penetration Adaptable aperture	Increased scattering artifacts (not fully correctable) Lower SNR than ATR Minimal sample thickness for sufficient absorption Maximal sample thickness limit for sufficient signal collection
ATR	Evanescent or transient wave is created by total internal reflection of incident wave within crystal (Internal Reflection Element, IRE) with high refractive index (e.g., Ge, Si, diamond…) in direct contact with sample. Internally reflected wave gets partially attenuated and detected.	Calcium or barium fluoride, zinc selenide, MirrIR low-E coated glass	Depends on refractive index of the IRE, aperture and diffraction limit of the microscope objective.[67, 68] (variable)	Only weak 'physical artifacts' due to non-complete contact with IRE (minor in comparison to scattering artifacts) High signal-to-noise ratio (SNR) No maximal sample thickness	Destructive Shallow sample penetration (~0.20 to 1.50 µm for organic compounds and 45° angle of incidence, depends on IRE material, wavelength: longer wavelengths penetrate deeper) Aperture only adaptable to discrete values by IRE choice

Table 4. Details on the three main FTIR spectroscopy sampling modes applied in this study. Adapted from Baker et al.[18] with respect to the measurement of individual dinocysts.

Supplementary figures

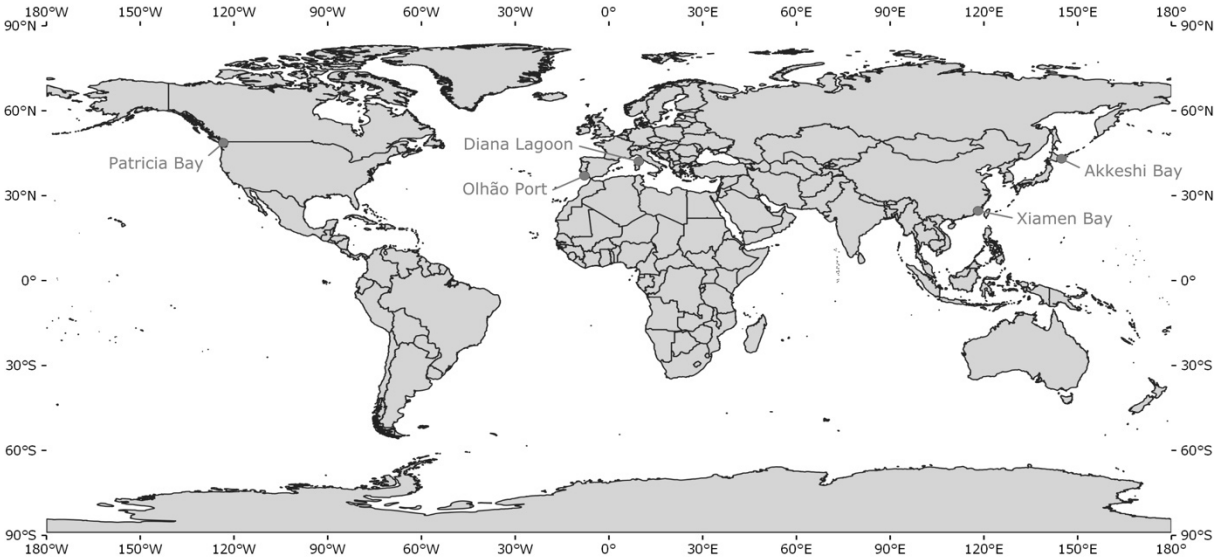


Figure S1. Map showing the locations of the samples analyzed in this study.

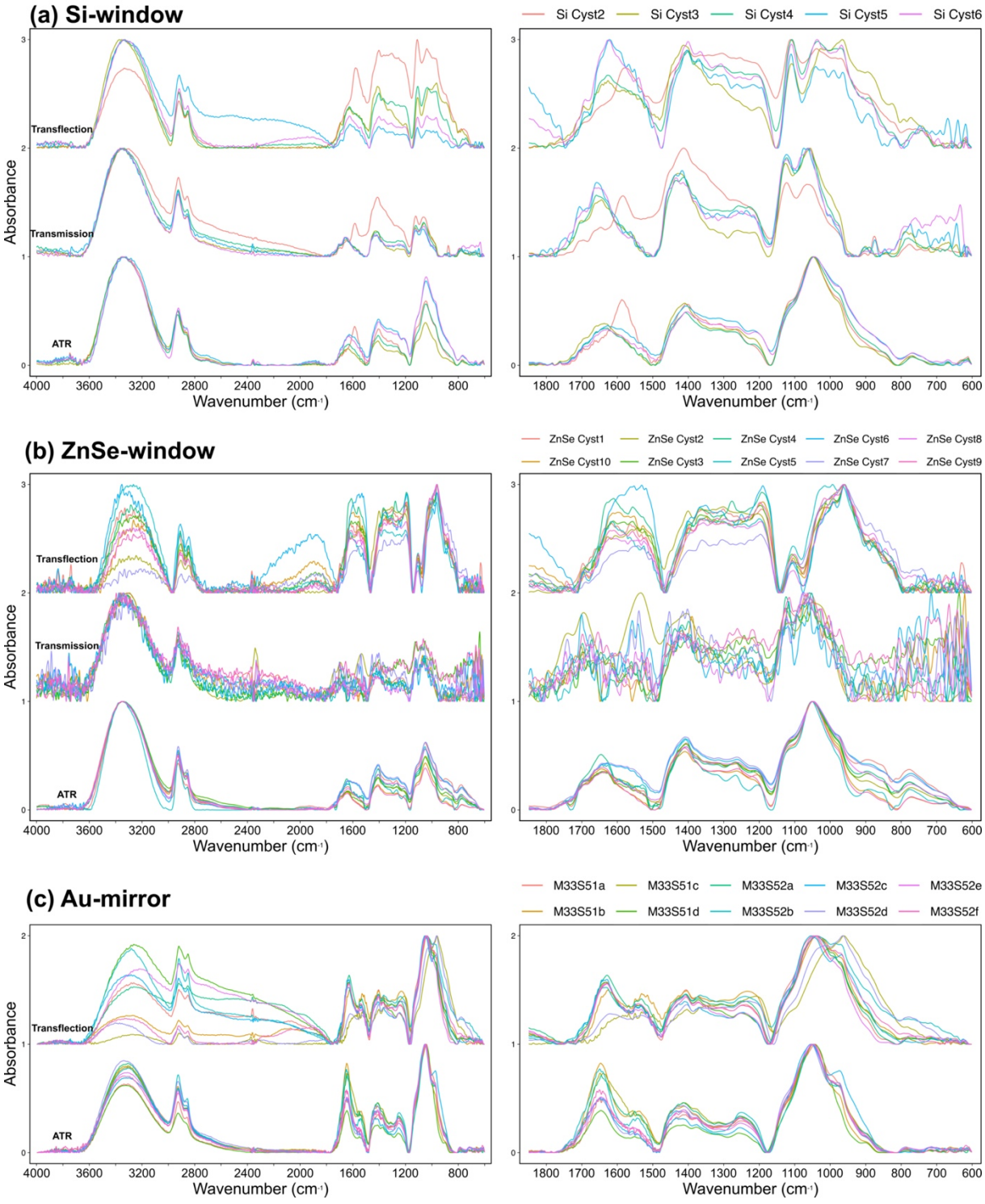


Figure S2**.** Individual measured culture-derived *Lingulodinium machaerophorum* spectra, using different sampling modes and measurement substrates (a–c). Full spectra (left) and fingerprint regions (right) both min-max normalized over their respective extent. Legends show original filenames of spectra. Note that Si Cyst2 strongly deviates from all others in all cases, strongly suggesting a different chemical composition recorded.

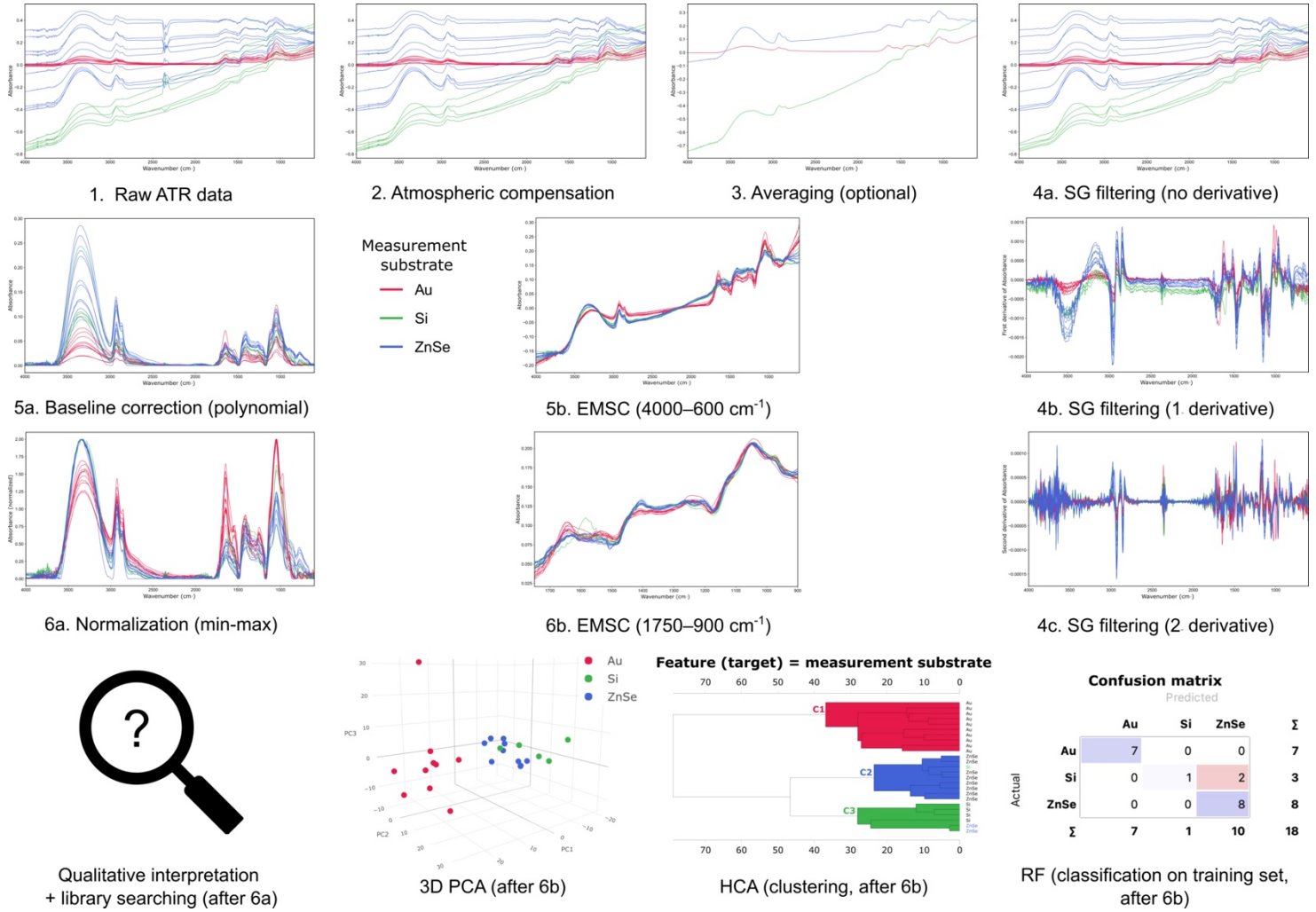


Figure S3. Graphic representation of a sequential data processing procedure on culture derived *Lingulodinium machaerophorum* ATR micro-FTIR data, recorded on different measurement substrates. Processing steps depend on the desired analysis goal. Abbreviations: SG = Savitzky-Golay, EMSC = extended multiplicative signal correction, PCA = principal component analysis, HCA = hierarchical cluster analysis, RF = random forest.