

: Tags >

[REPLACER](#)

alt. ver.: [Protein-Lipid Concentration Reference Measurements \(?\)](#)

Sample Prod.

- **rehydrate** a crumb of **product biomass** w/ 10x (by mass) water
- thick CaF₂ window on bottom
- steel spacer ring
- pipette 2µl rehydrated biomass into the middle
- close sandwich with thin CaF₂ window on top

Measuring

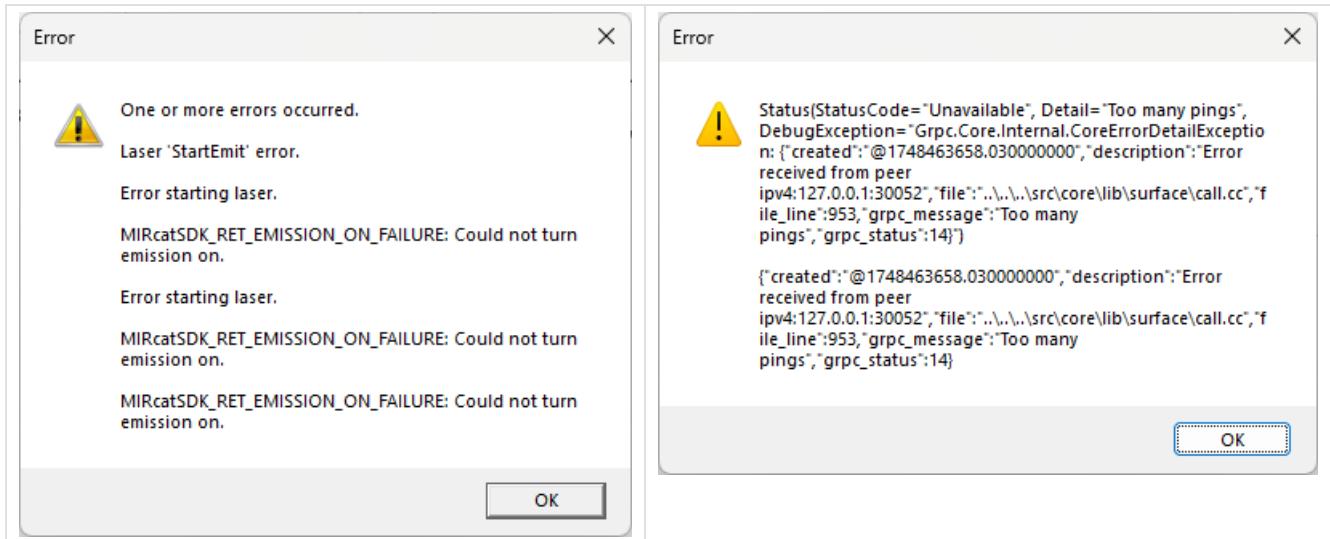
Copropagation, 40x Cassegrain

Auto-Background before 1st measurement

Averaging: 3

Sample	Labelled	Focus [µm]	Detector Gain
1	0,1476	2660.7	10x
2	illegible (cap blue, paper sticker: 0.01.25)	2484.5	50x
3	30.01.25 A. platensis 0,15 g	2649.9	50x
4	0,1653 g	2592.0	50x
5	98,6 mg	2657.2	50x
6	03.02.25 A. platensis 0,19 g	2569.8	50x
7	A. platensis	2641.9	50x
8	A. platensis 0,20 g 26.01.25	2578.8	50x
9	A. platensis 0,23 g		
10			
11			
12			

After production of the ninth sample, a crash of the setup was noticed:



The crash was unfortunately determined to be irrecoverable. No further measurements could be made.

Results

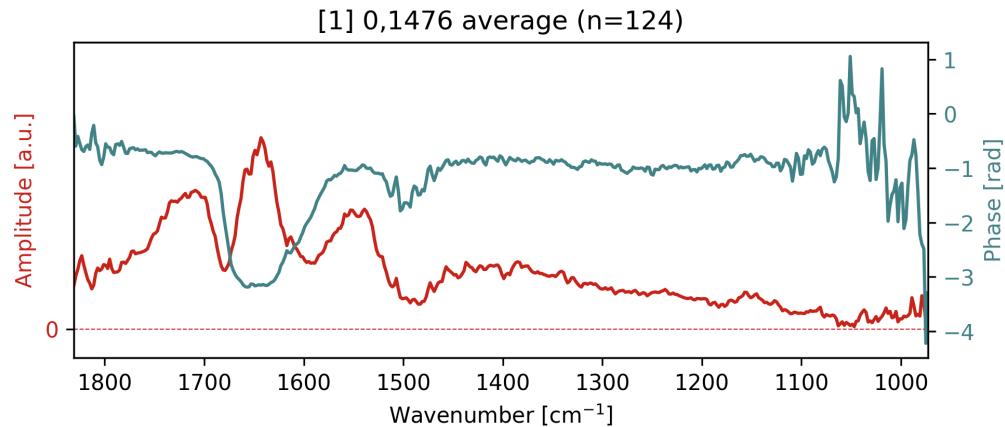
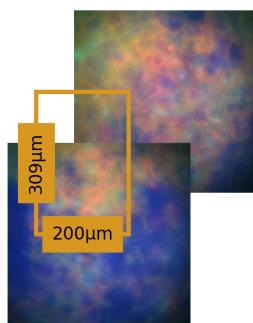
Inspection

Individual Spectra

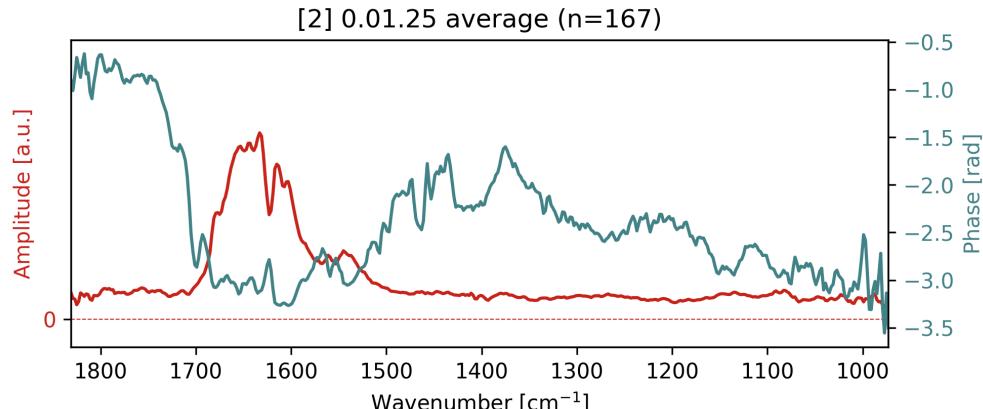
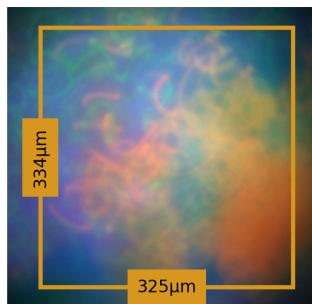
very noisy

Average Spectra

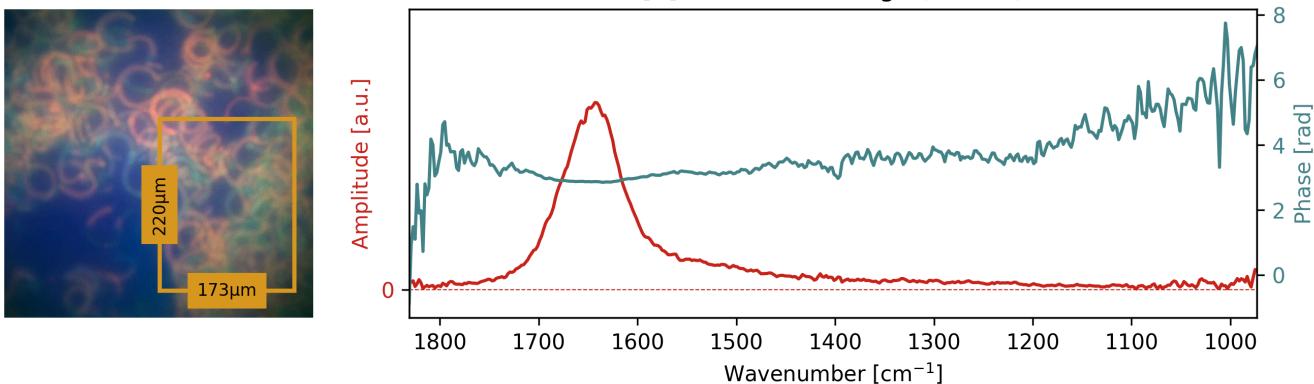
For each sample, an autofluorescence composite micrograph is given. The indicated areas are those, where spectra were recorded.



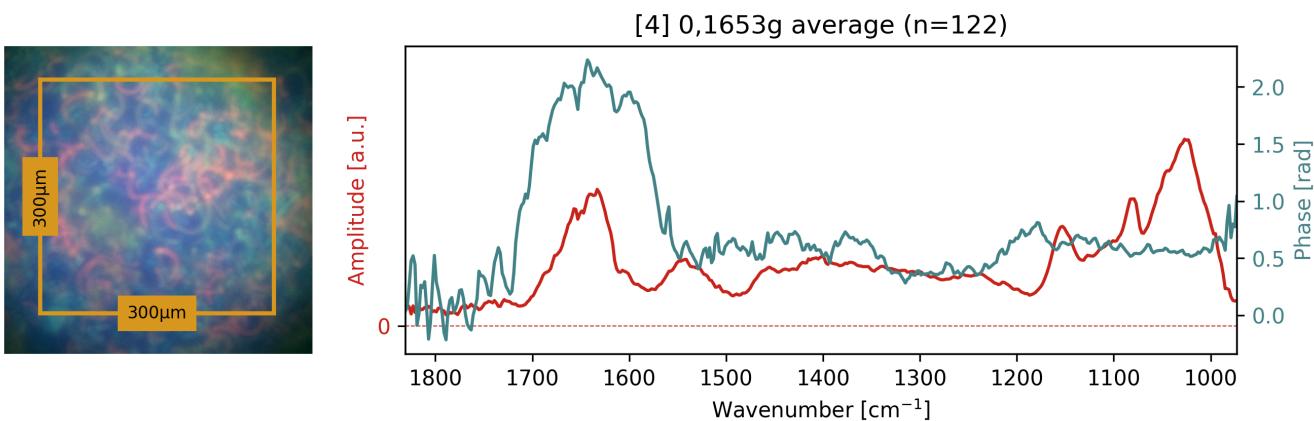
[1] Note how the main peak has a different phase from the secondary peaks. This might hint at the main peak corresponding to heating of a large volume (water background), taking a longer time for the temperature profile to relax and the secondary peaks corresponding to actual biomolecules, making up much less volume and therefore relaxing more quickly.



[2] Here, we similarly see a lagging phase in the region around 1650 cm⁻¹.

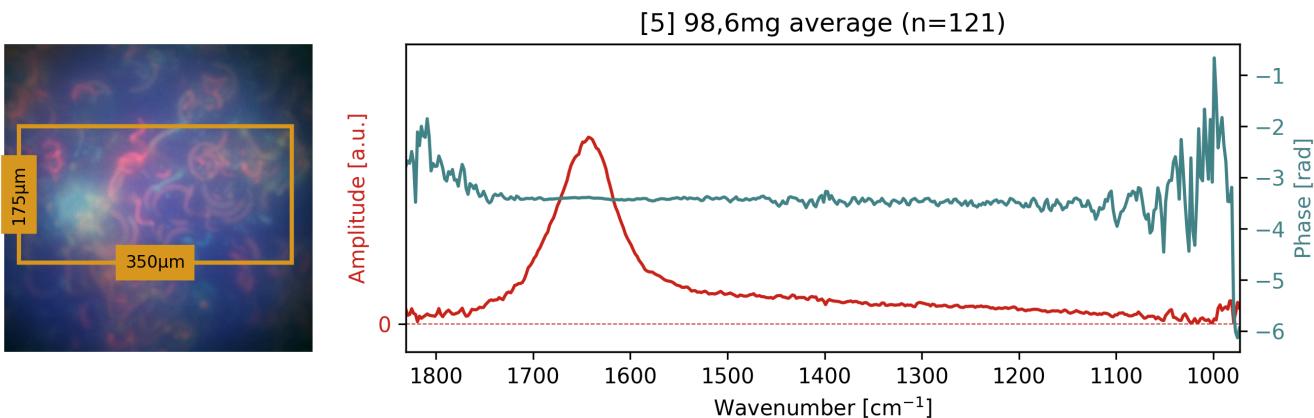


[3] Here, the phase seems to be invariant. The amplitude looks very smooth with only one visible peak and no secondaries. It doesn't look like any biomass is detected here, for some reason.

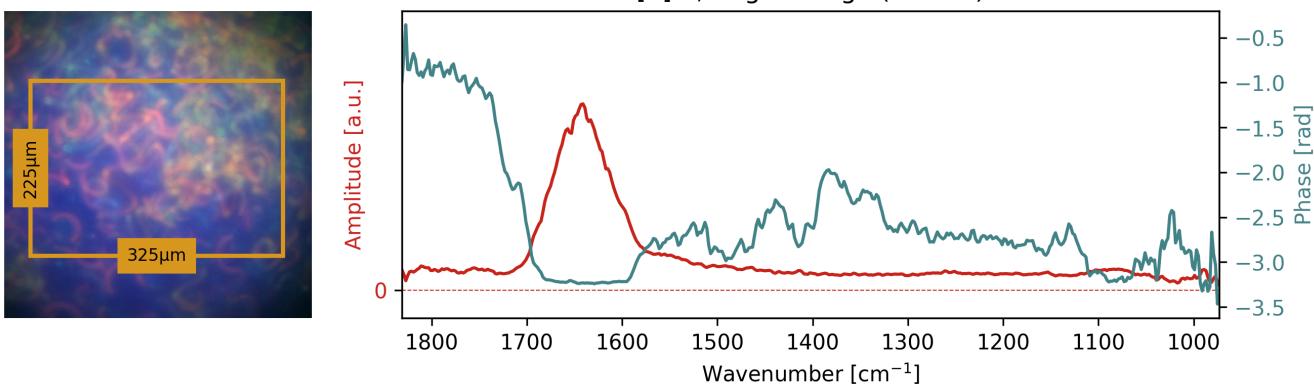


[4] Here, the phase doesn't lag around the water absorption but instead seems to advance. Reason unknown. Many additional peaks are visible in the amplitude of the signal. Their locations are consistent with usual bio-signatures, e.g.

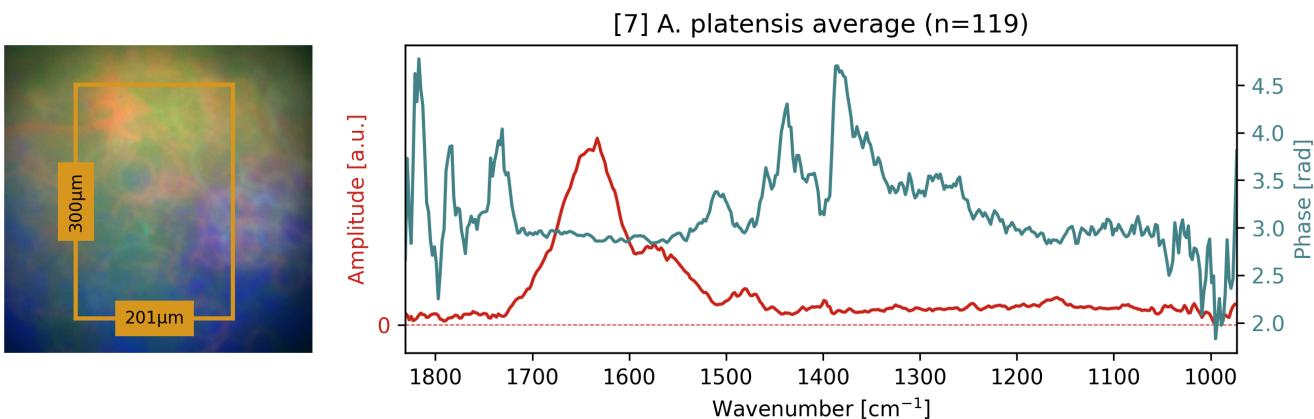
$\sim 1550 \text{ cm}^{-1}$	$\sim 1460 \text{ cm}^{-1}$	$\sim 1150 \text{ cm}^{-1}$	$\sim 1080 \text{ cm}^{-1}$	$\sim 1020 \text{ cm}^{-1}$
Amide II band	CH_2 scissor		PO_4 stretch	



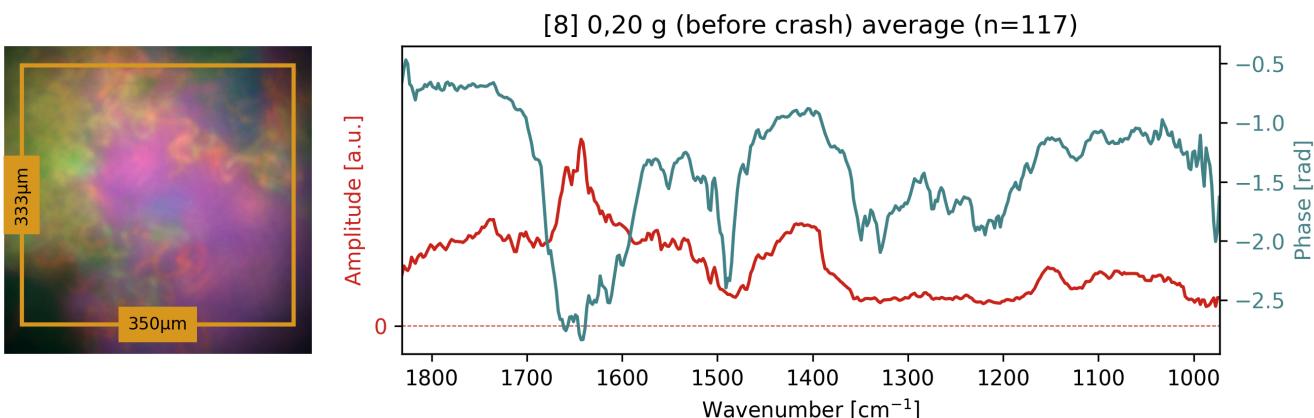
[5] as in [3], the phase is remarkably constant and no secondary peaks are visible.



[6] Secondary peaks are almost invisible. However, the phase clearly lags around the water absorption range.



[7] A peak in the high 1500s appears more like a shoulder to the side of the main peak at 1650 cm^{-1} . Smaller peaks are discernible in the amplitude at lower wavenumbers, and coincide with clear local phase lags.



[8] Multiple peaks with irregular shapes are discernible in the amplitude. Many features (moth peaks and valleys) of the amplitude spectrum coincide with features in the phase spectrum.

Analysis

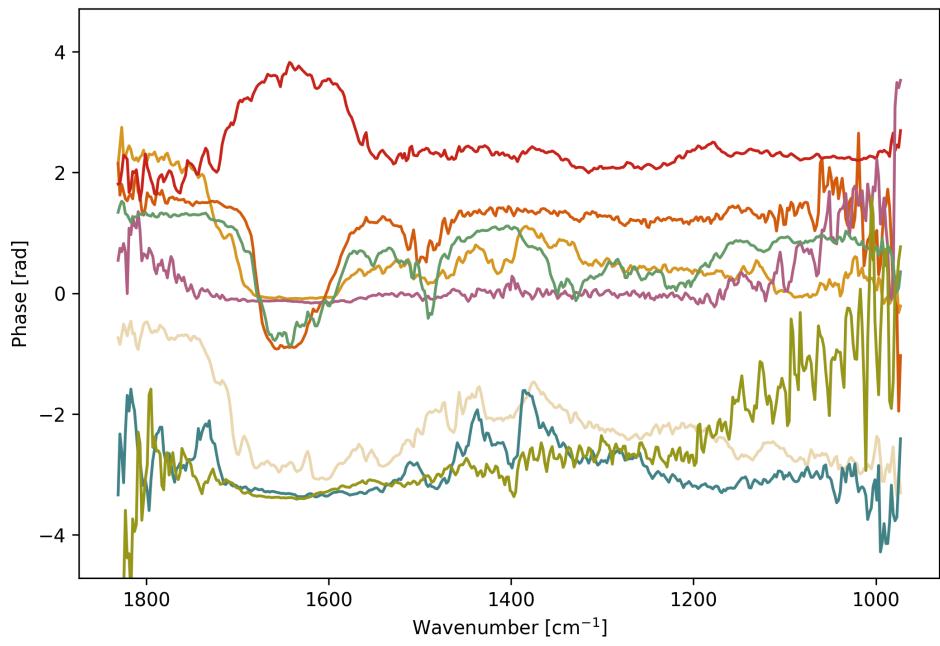
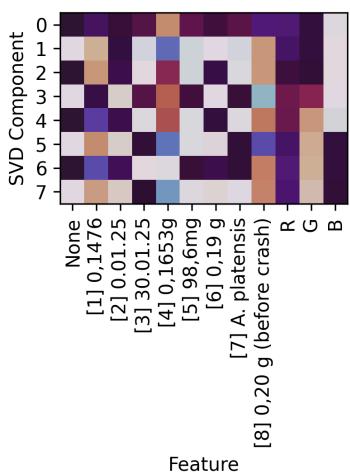
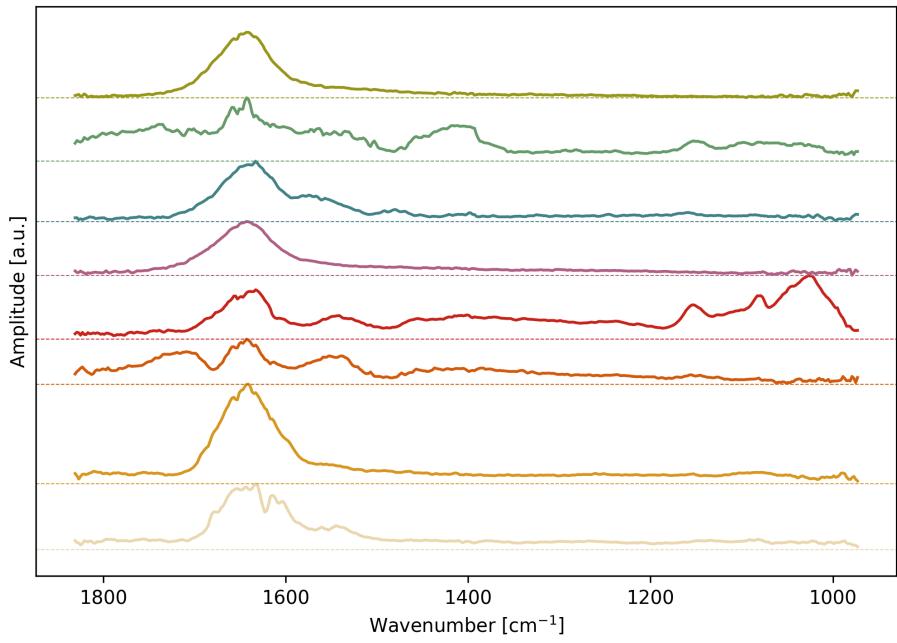
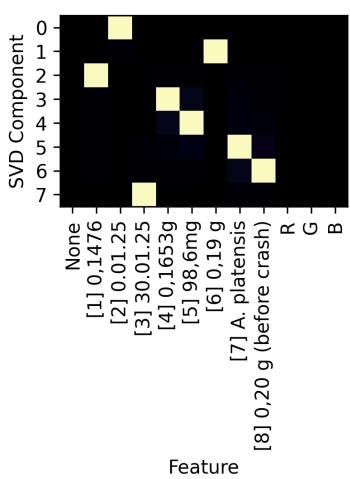
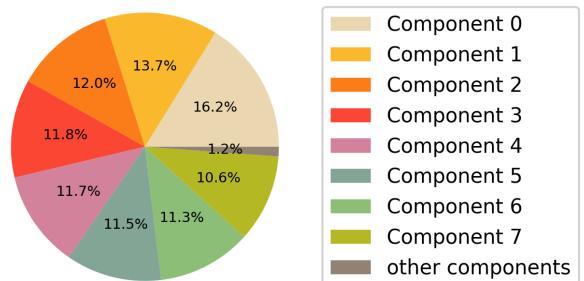
It appears prudent to analyse the complex spectra and not disregard the phase.

Singular Value Decomposition

I performed PCA (or SVD, because `sklearn.decompose.PCA` doesn't support complex-valued input) The explained variance per SVD component looked like so: (→)

This means that the first eight components were significant and the rest could be disregarded, essentially. Inspecting the components gave the following:

Variance Explained

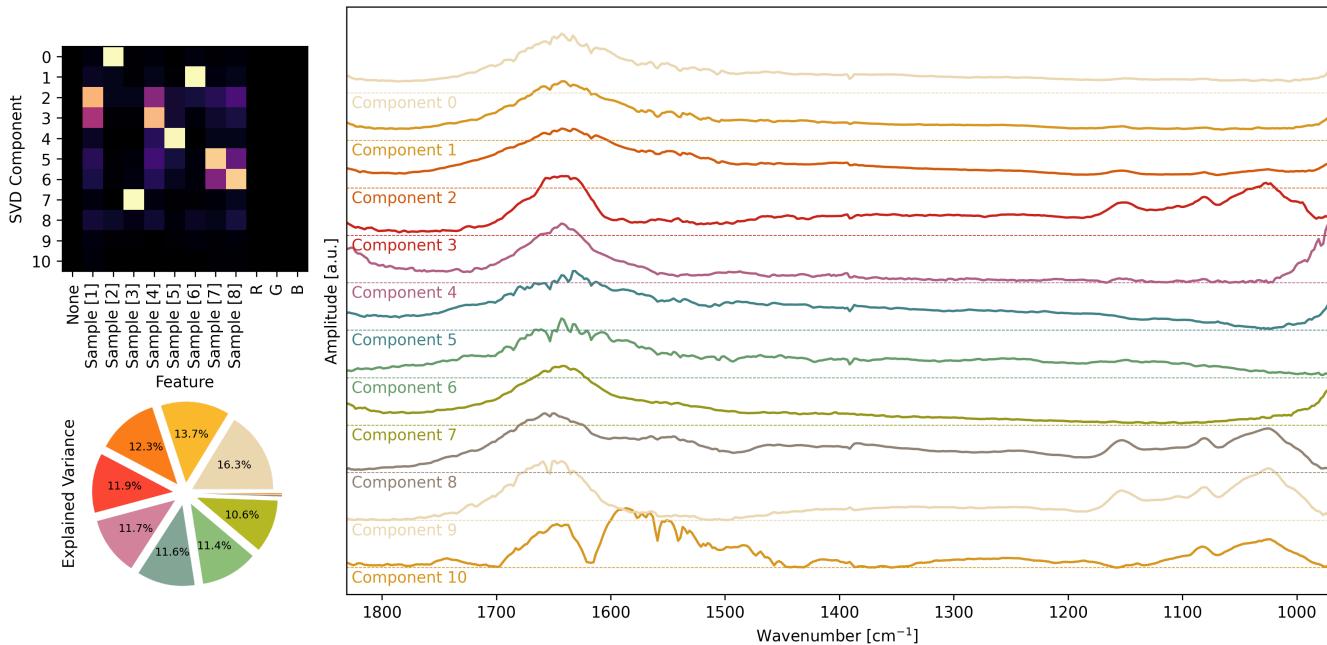


- **top left:** magnitude of the mapping between SVD components and features (sample indices)
- **top right:** magnitude of the SVD components vs wavenumbers
- **bottom left:** phase of the mapping between SVD components and features (sample indices)
- **bottom right:** phase of the SVD components vs wavenumbers

The interpretation is as straight-forward as it is disappointing:

The eight relevant SVD components were nothing but the average spectra of the eight samples.

When running the **SVD only on the amplitude** of the signal and disregarding the phase yields similar results:



- **top left:** magnitude of the mapping between SVD components and features (sample indices)
- **bottom left:** explained variance per SVD component. only 0.6% of the total variance of the dataset are explained by components beyond 7.
- **right:** SVD components vs wavenumbers

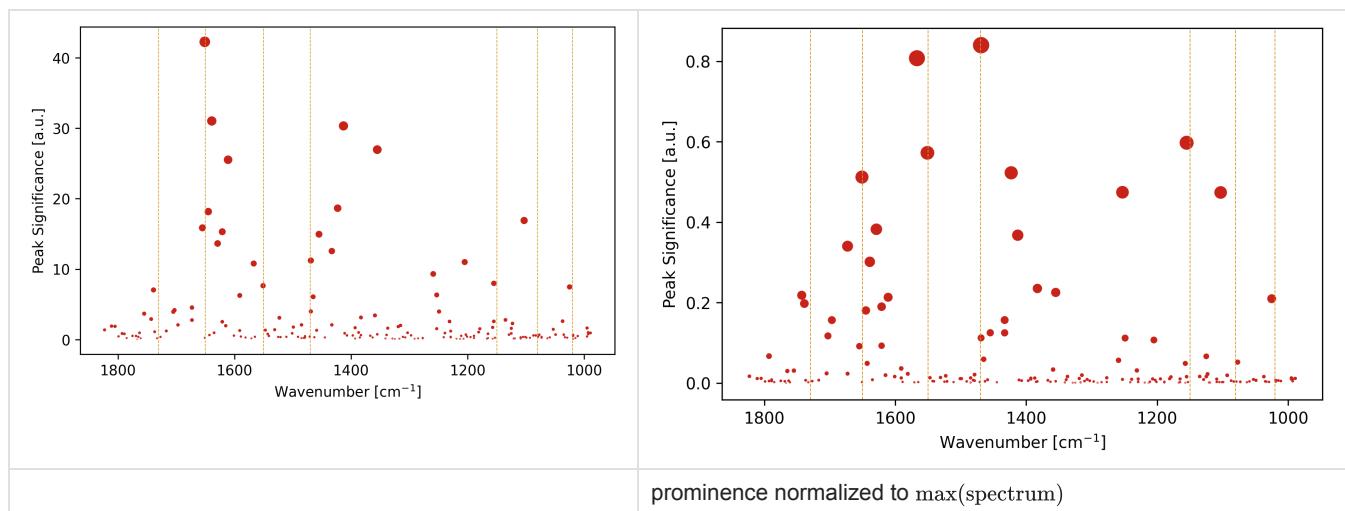
However, there is some overlap between measurements now. Still, it ought to be very clear that we should attempt to model the individual peaks and try decomposing into those.

Guided Decomposition

Let's accumulate information about all spectral peaks using `scipy.signal.find_peaks`. Obviously important is the position (in terms of wavenumbers) of each peak, but also, we require a measure for total absorption that a peak is associated with. Given the values of "prominence" and "width" that the `find_peaks` algorithm computes, we define a peak's

$$\text{significance} := \text{prominence} \cdot \text{width} .$$

Then, we can visualize all peaks like so:



The vertical lines highlight wavenumbers where we expect biosignatures.

Next, we require a way to sort these peaks into clusters.