Characterization and visualization of extracellular polymeric substances in anaerobic granular sludge

de Bruin, S.

DOI
10.4233/uuid:6c67edab-6686-419c-bfd6-4ad57ab56aee

Publication date
2024

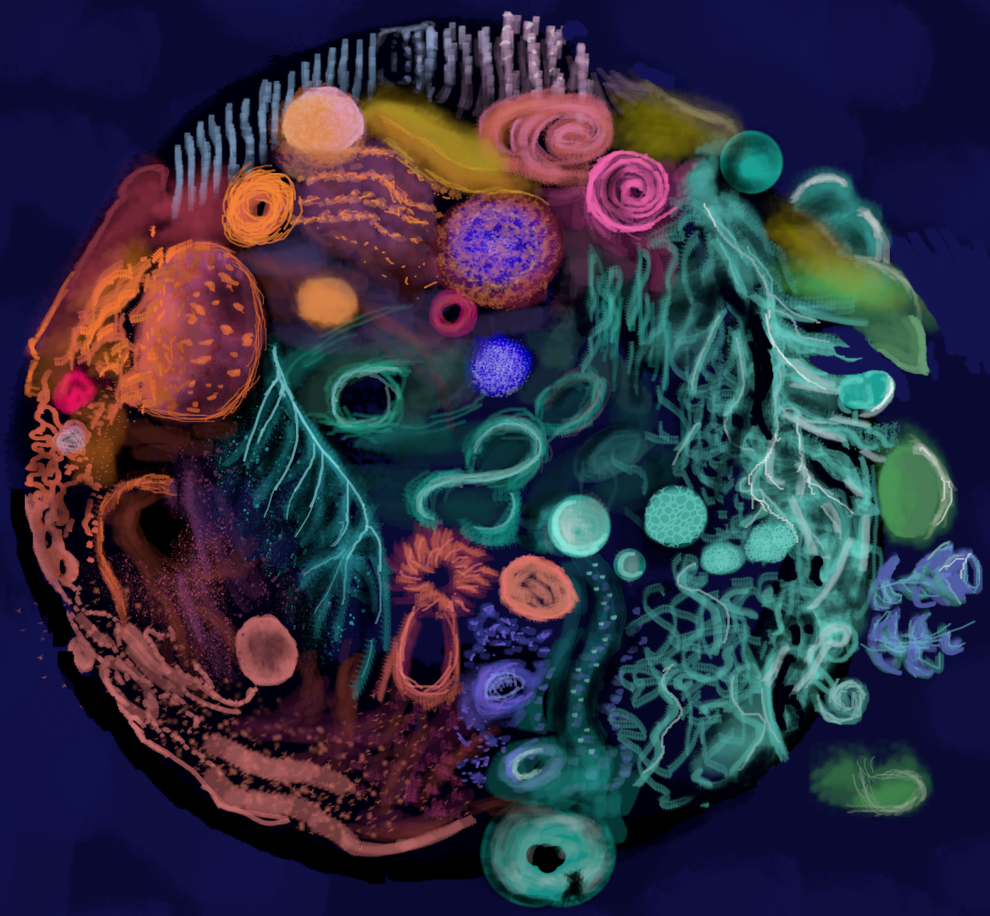
Document Version
Final published version

Citation (APA)
de Bruin, S. (2024). *Characterization and visualization of extracellular polymeric substances in anaerobic granular sludge*. [Dissertation (TU Delft), Delft University of Technology].
https://doi.org/10.4233/uuid:6c67edab-6686-419c-bfd6-4ad57ab56aee

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

S. de Bruin Characterization and visualization of extracellular polymeric substances in anaerobic granular sludge

Characterization and visualization of extracellular polymeric substances in anaerobic granular sludge



Stefan de Bruin

Propositions

Accompanying the dissertation
Characterization and Visualization of
Extracellular Polymeric Substances in Anaerobic Granular Sludge
by
Stefan DE BRUIN

1. The human-centric focus in medical microbiology biofilm research overlooks the ecological significance of sialic acids and sulfated glycans (**Chapters 2 & 3**)

2. Understanding EPS synthesis pathways should start by analyzing EPS composition rather than focusing on genomic potential (**Chapter 4**)

3. With a bit of creativity, any polymer can be found in the FT-IR spectrum of a biofilm; FT-IR analysis should always be supplemented with additional analyses (**this thesis**)

4. EPS research resembles piecing together a puzzle with missing parts, unexpected extras, and no initial idea of the final outcome (**this thesis**)

5. Interactions among diverse microorganisms are fundamental to the intricate dynamics inside a biofilm, this complexity cannot be replicated in a single-species biofilm

6. Exploratory research i.e. research without a predefined goal, warrants greater recognition and appreciation potential

7. Data interpretation and understanding should not be confused with advanced analysis methods and fancy figures (Forscher, B. K. (1963). Chaos in the brickyard. *Science*, *142*(3590), 339-339.)

8. Like collaborations inside biofilms, true altruism in society does not exist

9. Being able to simply explain your research should be trained more often during a PhD than only at the laymen's talk

10. If there is anything we can learn from Covid, it is how quickly a lifestyle can be enforced

These propositions are regarded as opposable and defendable, and have been approved as such by the promotor prof. dr. ir. M.C.M. Van Loosdrecht, promotor prof. dr. D.Z. Machado de Sousa and promotor dr. Y. Lin

Characterization and Visualization of Extracellular Polymeric Substances in Anaerobic Granular Sludge

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus, prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Thursday 19 September 2024 at 10:00 o'clock

by

Stefan DE BRUIN

Master of Science in Life Science and Technology,
Delft University of Technology, Netherlands
born in San José, Costa Rica

This dissertation has been approved by the

Promotor: prof. dr. ir. M.C.M. van Loosdrecht
Copromotor: prof. dr. D.Z. Machado de Sousa
Copromotor: dr. Y. Lin

Composition of the doctoral committee:

Rector Magnificus	chairperson
prof. dr. ir. M.C.M. van Loosdrecht	Delft University of Technology, promotor
prof. dr. D.Z. Machado de Sousa	Wageningen University and Research, promoter
dr. Y. Lin	Delft University of Technology, promotor

Independent members:

prof. dr. ir. C.J.N. Buisman	Wageningen University and Research
prof. dr. ir. J.B. van Lier	Delft University of Technology
dr. S. Roy	Delft University of Technology
prof. dr. M.C. Veiga-Barbazán	Universidade da Coruña, Spain
prof. dr. F. Hollmann	Delft University of Technology, reserve

The research presented in this thesis was performed at the Environmental Biotechnology Section, Department of Biotechnology, Faculty of Applied Sciences, Delft University of Technology, The Netherlands. The research was financially supported by the SIAM Gravitation Grant 024.002.002 from the Dutch Ministry of Education, Culture and Science and the Netherlands Organization for Scientific Research.

Table of contents

Summary

Wastewater from the food and agro-industry is filled with organic contaminants. If these substances are discharged into surface water, they promote the growth of unwanted microorganisms. To prevent this, contaminants are removed from the water through wastewater treatment. This is done using anaerobic digestion with microorganisms. Various types of microorganisms convert the organic compounds into methane gas, recovering some of the energy. The final step in anaerobic digestion, the conversion to methane, is the limiting factor in the process. A high concentration of methane-producing archaea is desired for rapid methane production.

Growing microorganisms in granules enables them to remain longer in the reactor, leading to an increased biomass concentration. The granules consist of multiple layers, each layer containing organisms that perform specific steps in the conversion to methane. These granules are a specific type of biofilm, made up of microorganisms embedded in a self-produced extracellular matrix. This matrix is composed of extracellular polymeric substances (EPS), which are produced and secreted by the microorganisms in the biofilm. EPS are a complex combination of proteins, polysaccharides, and lipids. Besides these basic polymers, combinations such as glycoproteins and lipopolysaccharides are also produced by the microorganisms. Charged polymers can form a polymer network with oppositely charged polymers or ions, contributing to the strength of the granular sludge. It is, therefore, no surprise that negatively charged particles, or acidic polymers, are often found in biofilms. However, how specific components in the EPS composition affect the structure and physical properties of granular sludge has been unclear until now.

The aim of this thesis is to study the EPS composition of anaerobic granular sludge, focusing on three main aspects: the identification of specific polymers, visualization of these polymers in the extracellular matrix, and identification of EPS synthesis pathways. While this thesis primarily characterizes the EPS composition of anaerobic granular sludge, the findings and methods are applicable to biofilms in general. By gaining a better understanding of the role of specific EPS components, we can better control biofilm processes.

In **Chapter 2**, we investigated negatively charged polymers in the EPS. Through chemical analysis, it was possible to measure the presence of negatively charged sulfated glycosaminoglycan-like polymers. Genome analysis of dominant microorganisms in the biofilm revealed that essential genes for the synthesis of sulfated polymers were present, suggesting that microorganisms can produce and not just accumulate sulfated polymers from their environment.

Chapter 3 examined the similarities in the molecular weight distribution pattern of EPS from granular sludge in aerobic and anaerobic wastewater treatment plants.

EPS was separated by gel permeation chromatography and showed a similar molecular weight distribution for both systems. The high-molecular-weight EPS fraction mainly consisted of glycoconjugates, with sialic acids predominantly found in this fraction. However, sulfated glycoconjugates were found in all molecular weight fractions. The similar distribution of molecular weight and negatively charged particles suggests that negatively charged particles perform similar functions in both biofilms.

In **Chapter 4**, we analyzed samples of anaerobic granular sludges, investigating the polysaccharide composition and using metagenome-assembled genomes (MAGs) to explore potential biosynthetic pathways. Uronic acids were the predominant components of the glycans in extracellular polymeric substances (EPS) produced by the anaerobic granular sludges, comprising up to 60% of the total polysaccharide content. MAGs associated with Anaerolineacae, Methanobacteriaceae, and Methanosaetaceae represented the majority of the microbial community (30-50% of total reads per MAG). Based on MAG analysis, it appeared that Anaerolinea sp. and members of the Methanobacteria class are involved in the production of exopolysaccharides within the analyzed granular sludges.

In **Chapter 5**, FT-IR microspectroscopy was used on slices of anaerobic granular sludge to investigate the EPS distribution. Visualization of specific functional groups in the absorption spectrum showed that the polysaccharide content was higher at the outer edge of the granule, while the protein concentration increased towards the center. This transition occurred approximately 150 μm from the outer edge, corresponding to literature descriptions of the size of the fermentative bacteria layer. Due to the complexity of the polymers, annotation was challenging. Therefore, principal component analysis and two-dimensional correlation spectroscopy were added to the analysis. These methods enabled the identification of overlapping functional groups and their correlations.

Finally, **Chapter 6** provides an overview of the key findings in this thesis and offers an outlook on directions for future research.

Samenvatting

Afvalwater van de voedsel- en agro-industrie is gevuld met organische contaminanten. Als deze stoffen in het oppervlakte water worden geloosd, dan zorgt dit voor de groei van ongewenste micro-organismes. Om dit tegen te gaan worden contaminanten uit het water gehaald door het afvalwater te behandelen. Dit gebeurt door middel van anaerobe vergisting met micro-organismes. Verschillende soorten micro-organismen zetten de organische verbindingen om tot methaan gas, waardoor een deel van de energie word teruggewonnen. De laatste stap in anaerobe vergisting, de omzetting tot methaan, is de beperkende factor in het proces. Voor een snelle methaanproductie is een hoge concentratie methaan-producerende archaea gewenst.

Door micro-organismen in korrels te laten groeien, kunnen ze langer in de reactor blijven, wat leidt tot een verhoogde biomassa concentratie. De korrels zijn opgebouwd uit meerdere lagen. Elke laag bevat organismen die een bepaalde stap van de omzetting naar methaan kunnen uitvoeren. Deze korrels zijn een specifiek soort biofilm en bestaan uit micro-organismen ingebed in een zelf geproduceerde extracellulaire matrix. Deze matrix is opgebouwd uit extracellulaire polymere substanties (EPS) die worden geproduceerd en uitgescheiden door de micro-organismen in het biofilm. EPS zijn een complexe combinatie van onder andere: eiwitten, polysacharide, en vetten. Naast deze basis polymeren, worden ook combinaties zoals bijvoorbeeld glycoproteïnes en lipo-polysacharides door de micro-organismen geproduceerd. Geladen polymeren kunnen een polymeer netwerk vormen met behulp van tegengestelde geladen polymeren of ionen en dragen daarom bij aan de stevigheid van het korrelslib. Het is daarom geen verrassing dat negatief geladen deeltjes, of wel zure polymeren, vaak worden aangetroffen in biofilms. Echter, hoe specifieke componenten inde EPS compositie de structuur en fysische eigenschappen korrelslib beïnvloeden was tot nu toe nog onduidelijk.

Het doel van dit proefschrift is om de EPS compositie van anaeroob korrelslib te bestuderen. Hierbij focussen we op drie hoofdpunten: De identificatie van specifieke polymeren, visualisatie van deze polymeren in de extracellulaire matrix en het identificeren van EPS synthese routes. Hoewel in dit proefschrift voornamelijk de EPS compositie van anaeroob korrelslib word gekarakteriseerd, zijn de bevindingen en methodes bruikbaar voor biofilms in het algemeen. Door een beter begrip te hebben van de rol van specifieke EPS componenten, kunnen we biofilm processen beter sturen.

In **hoofdstuk 2** deden we onderzoek naar negatief geladen polymeren in het EPS. Door middel van chemische analyse was het mogelijk om de aanwezigheid van negatief geladen gesulfateerde glycosaminoglycanen-achtige polymeren te meten. Uit de genoom analyse van dominante micro-organismen in het biofilm was het

duidelijk dat essentiële genen voor de synthese van gesulfateerde polymeren aanwezig waren. Dit suggereerde dat micro-organismen gesulfateerde polymeren kunnen produceren en niet alleen accumuleren uit hun omgeving.

Hoofdstuk 3 onderzocht de gelijkenissen in het molecuulgewichtspatroon van EPS uit korrelslib van aerobe en anaerobe afvalwaterzuiveringsinstallaties. EPS werd gescheiden door gelpermeatiechromatografie en toonde een vergelijkbare moleculaire gewichtsverdeling voor beide systemen. De hoogmoleculaire EPS-fractie bestond grotendeels uit glycoconjugaten, en siaalzuren werden vooral in deze fractie gevonden. Gesulfateerde glycoconjugaten echter, werden gevonden in alle molecuul gewicht fracties. De vergelijkbare verdeling van molecuulgewicht en negatief geladen deeltjes suggereert dat negatief geladen deeltjes vergelijkbare functies vervullen in beide biofilms.

In **hoofdstuk 4** analyseerden we monsters van anaerobe korrelslibben, waarbij we de polysacharide samenstelling onderzochten en metagenoom-geassembleerde genomen (MAGs) gebruikten om mogelijke biosynthese routes te verkennen. Uronzuren waren de belangrijkste bestanddelen van de glycanen in extracellulaire polymere stoffen (EPS) die werden geproduceerd door de anaerobe korrelslibben en vormden tot 60% van de totale polysacharide-inhoud. MAGs geassocieerd met Anaerolineacae, Methanobacteriaceae en Methanosaetaceae vertegenwoordigden de meerderheid van de microbiële gemeenschap (30-50% van de totale reads per MAG). Op basis van de analyse van MAGs bleek dat Anaerolinea sp. en leden van de Methanobacteria klasse betrokken zijn bij de productie van exo-polysachariden binnen de geanalyseerde korrelslibben.

In **hoofdstuk 5** werd FT-IR microspectroscopie gebruikt op plakken anaeroob korrelslib om de EPS-verdeling te onderzoeken. Visualisatie van specifieke functionele groepen in het absorptiespectrum liet zien dat de polysacharide-inhoud hoger was aan de buitenkant van de korrel, terwijl de eiwitconcentratie naar het centrum toenam. Deze overgang bevond zich ongeveer 150 μm van de buitenkant, wat overeenkomt met de literatuur over de grootte van de laag fermentatieve bacteriën. Vanwege de complexiteit van de polymeren was annotatie een uitdaging. Daarom werden hoofdcomponentenanalyse en tweedimensionale correlatiespectroscopie toegevoegd aan de analyse. Deze methoden maakten het mogelijk om overlappende functionele groepen en hun correlaties te identificeren.

Tot slot geeft **hoofdstuk 6** een overzicht van de belangrijkste bevindingen in dit proefschrift. Verder wordt een vooruitblik gegeven voor richtingen voor vervolg onderzoek.

Chapter 1:
General introduction

1.1. Anaerobic digestion and granular sludge

Industrial wastewaters rich in organic carbon are well-suited for anaerobic treatment by a consortium of anaerobic microorganisms. In this treatment, complex substrates present in the wastewater are ultimately converted into methane. Up to 1.5 kWh of energy can be recovered in the form of methane per kilo of removed chemical oxygen demand (COD) in organic compounds (Van Lier et al., 2015). This conversion of complex polymers to methane happens through a series of cascaded reactions, which are performed by different groups of microorganisms, as is showcased in figure 1.1. The process starts with hydrolysis, where complex polymers such as carbohydrates, lipids, and proteins are broken down into simpler soluble compounds. This step is facilitated by hydrolytic enzymes, which cleave large molecules into smaller fragments. In the acidogenesis step, the simpler compounds formed during the hydrolysis phase (e.g. simple sugars, amino acids, glycerol) are further metabolized by acidogenic bacteria, resulting in the production of volatile fatty acids (VFAs), alcohols, and other organic acids. Acetogenic bacteria turn the accumulated VFAs and organic acids into acetic acid, hydrogen, and carbon dioxide. This step is crucial as it prepares the substrate for the final stage of methanogenesis. The last step in the anaerobic digestion is the methanogenesis phase. Methanogenic archaea, convert the acetate into methane and carbon dioxide (aceticlastic methanogens) and H_2/CO_2 to methane (hydrogenotrophic methanogens) (Henze et al., 2008; Van Lier et al., 2015).

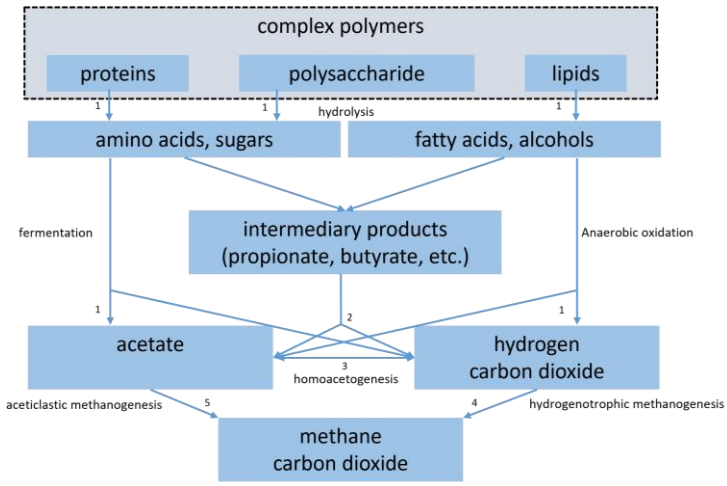


Figure 1.1. Overview of the catalytic cascade from complex polymers to methane. The reactions are numbered to show by which type of microorganism it is performed. 1. Hydrolytic and fermentative bacteria, 2. Acetogenic bacteria, 3. Homo-acetogenic bacteria, 4. Hydrogenotrophic archaea and 5. Aceticlastic archaea. Figure adapted from Henze et al. (2008).

The substrate uptake rate is related to the microbial growth rate, which is low for aceticlastic methanogens. High concentrations of biomass are needed to ensure high conversion rates inside the wastewater treatment systems. The biomass concentration increase is achieved by uncoupling the solid retention and hydraulic retention time with biomass retention strategies. This higher biomass concentration, results in high-rate treatment systems that can treat higher COD loads at relatively lower reactor volumes. Granular sludge reactors, like e.g. the Upflow Anaerobic Sludge Bed (UASB) or the Internal Circulation (IC) are prime examples of these high-rate reactors (Van Lier et al., 2015). Inside these reactors, microorganisms aggregate to form a granular macro-structure with diameter in the millimetre size range.

The microbial community inside the granule is not homogeneously distributed, but is instead divided into a layered structure with different microbial compositions. On the outer layer, closest to the surface, hydrolytic and fermentative microorganisms catabolize the complex polymers and convert them into VFAs. These are then consumed by the community in the next layer consisting of acetogenic bacteria and hydrogenotrophic methanogens. The core of the granule is usually populated by high numbers of aceticlastic methanogens (Macleod et al., 1990; Satoh et al., 2007; Doloman et al., 2017). Granulation mechanisms can be divided into categories being: physical, ecological and chemical (extensively

discussed in Hulshoff Pol et al., 2004). In UASB and EGSB reactors a high upflow velocity is applied which is suggested to trigger microbial aggregation. In these systems substrates are fed from the bottom of the reactor so fast-settling biomass reside in areas with a higher concentration gradient. Selective wasting of slow-settling biomass, because of the up flow in the reactor, further ensures a physical advantage for granule forming biomass. Ecologically, it is hypothesized that species like *Methanosaeta* are important for granulation as these form filamentous structures in the form of "spaghetti balls", which are later filled up with other microorganisms. It is true that *Methanosaeta* species are abundantly prevalent in anaerobic granular sludge, but other genera belonging to *Methanobacterium* and *Anaerolinea* genera have also been considered of importance for microbial aggregation (Xia et al., 2016; Zhu et al., 2017; Wu et al., 2021). Chemically, aggregation can be explained by the occurrence of excreted extracellular polymers. These polymers are found surrounding the microorganisms and aid in the adhesion and cohesion of the biofilm (Ras et al., 2013; Harimawan and Ting, 2016). However, how these extracellular polymers are influenced by microbial composition, substrate and other environmental factors is still largely unknown.

1.2. Extracellular polymeric substances in anaerobic granular sludge

A common agreement in biofilm studies is that the biofilm consists of a diverse community embedded within an extracellular matrix that is internally produced by the biofilm inhabitants. Important to note here is that granular sludge is considered a special type of spherical biofilm (Grotenhuis et al., 1991). Extracellular polymeric substances (EPS) are secreted by the inhabitants of the biofilm and are the building blocks of the extracellular matrix. The main components of these polymers are: proteins, polysaccharides, lipids and DNA (Flemming et al., 2023). In the biofilm, EPS protect microorganisms from e.g. shear, pressure and osmotic stress (Flemming et al., 2023). The EPS composition in anaerobic granular sludge are dominated by a high protein content in the extracellular matrix (Mills et al., 2023). Previous studies showed that granule core is highly enriched with proteins and the outside contains more polysaccharides (Morgan et al., 1991). Proteome analysis on the EPS proteins from anaerobic granular sludge revealed most identifiable proteins in the EPS were catalytic enzymes. Other abundant proteins in the EPS were S-layer proteins (Dubé and Guiot, 2019). S-layer proteins are crucial for biofilm creation and stabilization and are often glycosylated, which facilitates bonding between polymers and microorganisms (Sleytr et al., 2014; Pabst et al., 2022; Varki et al., 2015).

Charged species aid in the hydrogel formation capabilities of EPS, as these acidic polymers can cross-link with divalent cations. This justifies the enhanced granulation which occurred when divalent cations were added to an anaerobic reactor (Yu et al., 2001; Ahmad et al., 2011). It is therefore no surprise that acidic polymers are commonly found in all kinds of biofilms. Specifically, highly acidic polymers containing sulfate functional groups were found in various biofilm types (Bourven et al., 2015; Boleij et al., 2020; Felz et al., 2020). Other prevalent highly acidic polymers are polymers containing sialic acids groups. Sialic acids are a family of nine-carbon sugars with a large variability in their chemical structure. In animals, sialic acids are often found on the terminal end of glycan chains and are therefore important in the interaction between the glycan and the extracellular matrix. Fascinatingly, in bacterial biofilms, sialic acids seem to potentially play a similar function. Sialic acids were seen to be bound to galactose and protect it from galactosidase cleavage, thereby suggesting sialic acid being a protective terminal end group (De Graaff et al., 2019). In subsequent studies the widespread prevalence of sialic acids, sulfated and other acidic polymers was found in different biofilms (Boleij et al., 2020; Felz et al., 2020; Tomás-Martínez et al., 2021; Tomás-Martínez et al., 2022).

1.3. Characterization methods for extracellular polymeric substances

One of the challenges in EPS research is characterizing the enormous diversity of polymers present in the extracellular matrix. Due to the largely unexplored diversity of polymers, EPS has been coined 'the dark matter of biofilms'. Interestingly, dark matter refers to the invisible and largely undetectable mass in our universe. Of course the term is not completely accurate as the reason dark matter is difficult to detect is because it does not interact with light (De Swart et al., 2017). The problem of EPS is that actually too many functional groups interact in the measurements and so the signals of specific components are often convoluted through interference of other similar functional groups(Felz et al., 2019; Ali et al., 2018). The complexity of the EPS polymers is heightened by the presence of combinations of polymer types such as e.g. glyco-proteins and lipo-polysaccharides(Flemming et al., 2023).

Seviour et al. (2019) proposed a roadmap for understanding the role of individual EPS components in the extracellular matrix. Which consists of three main elements: identification of specific polymers, visualization of these specific components and identifying the synthesis pathways in genome data. The first obstacle for the identification of specific EPS polymers, is to extract it out of the extracellular matrix. Structural polymers were extracted out of granules by

performing alkaline EPS extraction which breaks down the biofilm structure (Felz et al., 2016). After extraction, general polymer content is often determined by chemical (colorimetric) methods, and through overlapping methods, effects from biases and interference from components in the sample are minimized (Felz et al., 2020). Isolation of specific EPS polymers is needed for the determination of the polymer structure. This is important as polymers are very similar in biological samples, which make vibrational spectroscopy and mass spectrometry attempts highly convoluted (Ali et al., 2018). Isolation has been performed in the past by using chemical assays targeting highly acidic species, like e.g. sulfated glycosaminoglycan-like polymers (Felz et al., 2020; Boleij et al., 2020). Additionally, separation based on gel chromatography (size) or cation exchange chromatography (charge) are also possible methods (Bourven et al., 2015; Gonzalez-Gil et al., 2015). Known polymer structures in pure culture biofilm studies aided in the navigation of the microbial community genomes concerning EPS synthesis (Dueholm et al., 2023). Despite these revelations, the synthesis pathways of EPS still remain poorly characterized. Which might be caused by the enormous variability of polymers. Two imaging routes for visualizing the spatial distribution of specific polymers have been suggested. The first are targeted techniques like e.g. antibody probe imaging, or lectin staining analysis. The second route comprises untargeted imaging techniques such as MALDI-TOF MS, Raman micro-spectroscopy, and FT-IR micro-spectroscopy (Seviour et al., 2019). Notably, FT-IR micro-spectroscopy, is known for its user-friendly and non-destructive nature. Using this equipment could therefore offer valuable insights into polymer distribution in the extracellular matrix (Gowen et al., 2015).

Therefore, given the complexity of this extracellular matrix and its multifaceted composition, novel measurement and data analysis approaches are needed for comprehensive analysis. Since numerous factors influence the composition of EPS, a robust exploratory methodology is crucial.

Thesis outline and scope

This thesis aims to elucidate the role of specific EPS components within the extracellular matrix of anaerobic granular sludge. It should be noted that the approaches herewith outlined are also suitable for the EPS analysis of other biofilms. A broader study of EPS in various biofilms would enhance our understanding of EPS physico-chemical characteristics, mechanisms of formation and funtions. This is why, EPS samples from various granular sludges and other biofilms were extracted and subjected to chemical analysis.

Essential in EPS research is to: accurately identify the types of polymers present in the extracellular matrix (chapter 2 and 3), finding the organisms responsible for producing these compounds (chapter 4), along with their localization in the biofilm (chapter 5). A schematic representation can be found in figure 1.2. By addressing these fundamental questions—what polymers are present and what functions they serve, who produces them, and where they are produced—this thesis embarks on an exploratory journey to characterize the extracellular matrix.

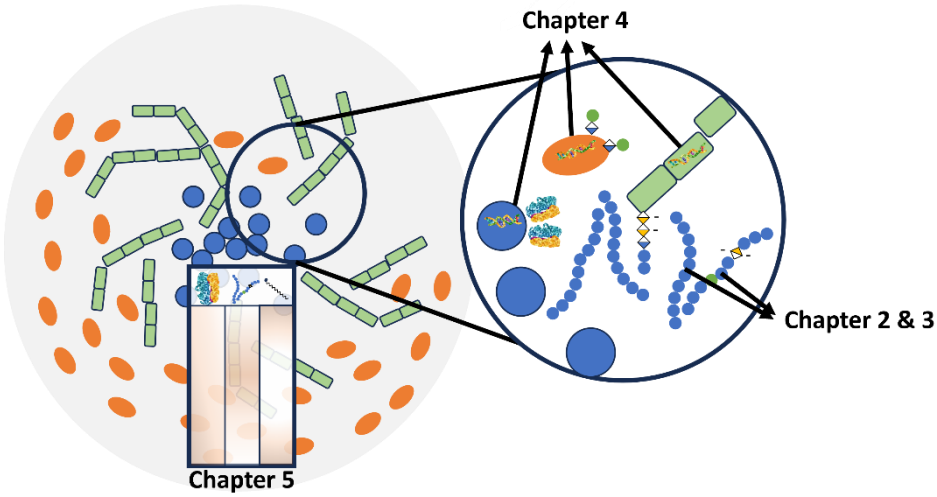


Figure 1.2. Schematic cross-cut overview of an anaerobic granule showing the focus points of the different chapters.

In **chapter 2** the presence of acidic polymers, namely sulfated glycosaminoglycan-like polymers are investigated. EPS was extracted from samples taken from a sulfur cave biofilm. The microorganisms inside the biofilm had no access to any organic carbon source. Therefore, if sulfated polymers are present, it shows they can be produced by the microorganism in the biofilm.

Chapter 3 shows how the sialic acids and sulfated polymers are distributed across polymers of different molecular weights. EPS from granular sludge samples from aerobic and anaerobic systems were separated based on molecular weight using size-exclusion chromatography. This was done to see if similar patterns in the anionic polymer distribution could be distinguished in the EPS of these systems.

In **Chapter 4**, an approach is proposed to utilize chemical data for guiding the exploration of genes involved in EPS synthesis pathways, which remain largely

unknown. EPS samples extracted from anaerobic granular sludge underwent analysis using a combination of chemical methods to measure the composition of polysaccharide sugar monomers. Subsequently, the metagenome of these samples was analyzed to target genes necessary for the biosynthesis of these sugar monomer precursors. This approach enabled the identification of microorganisms within the anaerobic granules with the potential to produce EPS.

Chapter 5 is focussed on the application of FT-IR imaging for visualizing the chemical composition of EPS throughout the granular sludge. This study involves mapping the relative concentrations of specific functional groups across slices of granular sludge. FT-IR micro-spectroscopy is an exploratory untargeted and non-destructive technique. The complex nature of the extracellular matrix makes annotating the spectra challenging. Through the use of chemometrics however, a more guided effort to visualize the polymer composition in the extracellular matrix can be achieved. The use of FT-IR micro-spectroscopy may enhance our understanding of the spatial arrangement of polymers in the granular sludge matrix, providing valuable insights into its composition and structure.

In **Chapter 6** a summary of the main findings in this thesis is provided, along with an outlook for future research directions informed by these findings.

References

Ahmad, A., Ghufran, R., & Wahid, Z. A. (2011). Role of calcium oxide in sludge granulation and methanogenesis for the treatment of palm oil mill effluent using UASB reactor. *Journal of hazardous materials*, *198*, 40-48.

Ali, M. H., Rakib, F., Al-Saad, K., Al-Saady, R., Lyng, F. M., & Goormaghtigh, E. (2018). A simple model for cell type recognition using 2D-correlation analysis of FTIR images from breast cancer tissue. Journal of Molecular Structure, 1163, 472-479.

Boleij, M., Kleikamp, H., Pabst, M., Neu, T. R., van Loosdrecht, M. C., & Lin, Y. (2020). Decorating the anammox house: sialic acids and sulfated glycosaminoglycans in the extracellular polymeric substances of anammox granular sludge. Environmental science & technology, 54(8), 5218-5226.

Bourven, I., Bachellerie, G., Costa, G., & Guibaud, G. (2015). Evidence of glycoproteins and sulphated proteoglycan-like presence in extracellular polymeric substance from anaerobic granular sludge. Environmental technology, 36(19), 2428-2435.

de Graaff, D. R., Felz, S., Neu, T. R., Pronk, M., van Loosdrecht, M. C., & Lin, Y. (2019). Sialic acids in the extracellular polymeric substances of seawater-adapted aerobic granular sludge. Water research, 155, 343-351.

De Swart, J. G., Bertone, G., & van Dongen, J. (2017). How dark matter came to matter. Nature Astronomy, 1(3), 0059.

Doloman, A., Varghese, H., Miller, C. D., & Flann, N. S. (2017). Modeling de novo granulation of anaerobic sludge. BMC systems biology, 11, 1-12.

Dubé, C. D., & Guiot, S. R. (2019). Characterization of the protein fraction of the extracellular polymeric substances of three anaerobic granular sludges. *AMB Express*, *9*, 1-11.

Dueholm, M. K. D., Besteman, M., Zeuner, E. J., Riisgaard-Jensen, M., Nielsen, M. E., Vestergaard, S. Z., ... & Nielsen, P. H. (2023). Genetic potential for exopolysaccharide synthesis in activated sludge bacteria uncovered by genome-resolved metagenomics. Water Research, 229, 119485.

Felz, S., Al-Zuhairy, S., Aarstad, O. A., van Loosdrecht, M. C., & Lin, Y. M. (2016). Extraction of structural extracellular polymeric substances from aerobic granular sludge. JoVE (Journal of Visualized Experiments), (115), e54534.

Felz, S., Neu, T. R., van Loosdrecht, M. C., & Lin, Y. (2020). Aerobic granular sludge contains Hyaluronic acid-like and sulfated glycosaminoglycans-like polymers. Water research, 169, 115291.

Felz, S., Vermeulen, P., van Loosdrecht, M. C., & Lin, Y. M. (2019). Chemical characterization methods for the analysis of structural extracellular polymeric substances (EPS). Water Research, 157, 201-208.

Flemming, H. C., van Hullebusch, E. D., Neu, T. R., Nielsen, P. H., Seviour, T., Stoodley, P., ... & Wuertz, S. (2023). The biofilm matrix: Multitasking in a shared space. Nature Reviews Microbiology, 21(2), 70-86.

Gonzalez-Gil, G., Thomas, L., Emwas, A. H., Lens, P. N., & Saikaly, P. E. (2015). NMR and MALDI-TOF MS based characterization of exopolysaccharides in anaerobic microbial aggregates from full-scale reactors. Scientific Reports, 5(1), 14316.

Gowen, A. A., Feng, Y., Gaston, E., & Valdramidis, V. (2015). Recent applications of hyperspectral imaging in microbiology. Talanta, 137, 43-54.

Grotenhuis, J. T. C., Smit, M., Van Lammeren, A. A. M., Stams, A. J. M., & Zehnder, A. J. B. (1991). Localization and quantification of extracellular polymers in methanogenic granular sludge. Applied microbiology and biotechnology, 36, 115-119.

Harimawan, A., & Ting, Y. P. (2016). Investigation of extracellular polymeric substances (EPS) properties of P. aeruginosa and B. subtilis and their role in bacterial adhesion. Colloids and Surfaces B: Biointerfaces, 146, 459-467.

Henze, M., van Loosdrecht, M. C., Ekama, G. A., & Brdjanovic, D. (Eds.). (2008). *Biological wastewater treatment*. IWA publishing.

MacLeod, F. A., Guiot, S. R., & Costerton, J. (1990). Layered structure of bacterial aggregates produced in an upflow anaerobic sludge bed and filter reactor. *Applied and environmental microbiology*, *56*(6), 1598-1607.

Mills, S., Trego, A. C., Prevedello, M., De Vrieze, J., O'Flaherty, V., Lens, P. N., & Collins, G. (2023). Unifying concepts in methanogenic, aerobic, and anammox sludge granulation. Environmental Science and Ecotechnology, 100310.

Morgan, J. W., Evison, L. M., & Forster, C. F. (1991). The internal architecture of anaerobic sludge granules. Journal of Chemical Technology & Biotechnology, 50(2), 211-226.

Pabst, M., Grouzdev, D. S., Lawson, C. E., Kleikamp, H. B., de Ram, C., Louwen, R., ... & Laureni, M. (2022). A general approach to explore prokaryotic protein glycosylation reveals the unique surface layer modulation of an anammox bacterium. The ISME Journal, 16(2), 346-357.

Pol, L. H., de Castro Lopes, S. I., Lettinga, G., & Lens, P. N. L. (2004). Anaerobic sludge granulation. Water Research, 38(6), 1376-1389.

Ras, M., Lefebvre, D., Derlon, N., Hamelin, J., Bernet, N., Paul, E., & Girbal-Neuhauser, E. (2013). Distribution and hydrophobic properties of extracellular polymeric substances in biofilms in relation towards cohesion. Journal of biotechnology, 165(2), 85-92.

Satoh, H., Miura, Y., Tsushima, I., & Okabe, S. (2007). Layered structure of bacterial and archaeal communities and their in situ activities in anaerobic granules. *Applied and environmental microbiology*, *73*(22), 7300-7307.

Seviour, T., Derlon, N., Dueholm, M. S., Flemming, H. C., Girbal-Neuhauser, E., Horn, H., ... & Lin, Y. (2019). Extracellular polymeric substances of biofilms: Suffering from an identity crisis. Water research, 151, 1-7.

Sleytr, U. B., Schuster, B., Egelseer, E. M., & Pum, D. (2014). S-layers: principles and applications. FEMS microbiology reviews, 38(5), 823-864.

Tomás-Martínez, S., Chen, L. M., Pabst, M., Weissbrodt, D. G., van Loosdrecht, M. C., & Lin, Y. (2023). Enrichment and application of extracellular nonulosonic acids containing polymers of

Accumulibacter. Applied Microbiology and Biotechnology, 107(2-3), 931-941.

Tomás-Martínez, S., Kleikamp, H. B., Neu, T. R., Pabst, M., Weissbrodt, D. G., van Loosdrecht, M. C., & Lin, Y. (2021). Production of nonulosonic acids in the extracellular polymeric substances of "Candidatus Accumulibacter phosphatis". Applied Microbiology and Biotechnology, 105, 3327-3338.

Van Lier, J. B., Van der Zee, F. P., Frijters, C. T. M. J., & Ersahin, M. E. (2015). Celebrating 40 years anaerobic sludge bed reactors for industrial wastewater treatment. *Reviews in Environmental Science and Bio/Technology*, *14*, 681-702.

Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., ... & Seeberger, P. H. (2015). Essentials of Glycobiology [internet].

Wu, J., Jiang, B., Kong, Z., Yang, C., Li, L., Feng, B., ... & Li, Y. Y. (2021). Improved stability of up-flow anaerobic sludge blanket reactor treating starch wastewater by pre-acidification: Impact on microbial community and metabolic dynamics. *Bioresource Technology*, *326*, 124781.

Xia, Y., Wang, Y., Wang, Y., Chin, F. Y., & Zhang, T. (2016). Cellular adhesiveness and cellulolytic capacity in Anaerolineae revealed by omics-based genome interpretation. *Biotechnology for Biofuels*, *9*(1), 1-13.

Yu, H. Q., Tay, J. H., & Fang, H. H. (2001). The roles of calcium in sludge granulation during UASB reactor start-up. *Water research*, *35*(4), 1052-1060.

Zhu, X., Kougias, P. G., Treu, L., Campanaro, S., & Angelidaki, I. (2017). Microbial community changes in methanogenic granules during the transition from mesophilic to thermophilic conditions. *Applied microbiology and biotechnology*, *101*, 1313-1322.

Chapter 2:
Sulfated Glycosaminoglycan-like Polymers are Present in an Acidophilic Biofilm from a Sulfidic Cave

Abstract

Sulfated glycosaminoglycans (sGAG) are negatively charged extracellular polymeric substances that occur in biofilms from various environments. Yet, it remains unclear whether these polymers are acquired from the external environment or produced by microbes in the biofilm. To resolve this, we analyzed the presence of sGAGs in samples of an acidophilic biofilm collected from Sulfur Cave in Puturosu Mountain (Romania), an environment that is largely inaccessible to contamination. A maximum of 55.16 ± 2.06 μg sGAG-like polymers were recovered per mg of EPS. Enzymatic treatment with chondroitinase ABC resulted in a decrease of the mass of the polymers, suggesting the structure of the recovered sGAG is similar to chondroitin. Subsequent FT-IR analysis of these polymers revealed absorbance bands at 1212 cm^{-1}, 1167 cm^{-1} and 900 cm^{-1}, indicating a possible presence of polysaccharides and sulfate. Analysis of genomic sequences closely related to those predominant in the acidophilic biofilm, contained genes coding for sulfotransferase (an enzyme needed for the production of sGAG), which supports the hypothesis of microbial synthesis of sGAGs within the biofilm.

Keywords

EPS; Extremophiles; Sulfur Cave; Romania

Highlights

- Biofilm from Sulfur Cave is a unique environment to study sulfated polymers.
- Sulfated glycosaminoglycan-like polymers were present in Sulfur Cave biofilm.
- Sulfated polymers might be produced internally by biofilm inhabitants.

Published as:

De Bruin, S., Vasquez-Cardenas, D., Sarbu, S. M., Meysman, F. J. R., Sousa, D. Z., van Loosdrecht, M. C. M., & Lin, Y. (2022). Sulfated Glycosaminoglycan-like Polymers are Present in an Acidophilic Biofilm from a Sulfidic Cave. Science of the Total Environment, 829, 154472.

1. Introduction

Biofilms are aggregations of microorganisms that are immobilized in a self-produced matrix composed of extracellular polymeric substances (EPS). EPS provide mechanical stability and scaffolds for microbial growth and protect microbes in the biofilm from *e.g.* desiccation, biocides, heavy metals, ultraviolet radiation, protozoan grazers (Flemming and Wingender, 2010). Negatively charged polymers, such as sulfated glycosaminoglycans (sGAG), are particularly important for increased water retention, due to their ability to form hydrogels by cross-linking with cations (Zykwinska et al., 2019). These glycans are large polysaccharide chains built from disaccharide blocks of an amino sugar and either uronic acid or hexose sugar. Sulfation of hydroxyl groups in the disaccharide building block makes the sGAGs amongst the most negatively charged polymers in nature (Köwitsch et al., 2017). The high negative charge make sGAGs have diverse bioactivities in multicellular organisms, and sGAGs like heparin are used pharmaceutically as an anticoagulant. The chemical production of sGAGs is difficult and time consuming and therefore sGAGs are mainly sourced from animal tissue (Da Costa et al., 2017).

In the past, microorganisms were thought to be incapable of producing sGAG-like polymers (Schiraldi et al., 2010). Yet, recent studies suggest that sGAG-like polymers are present in a wide variety of microbial biofilms (Boleij et al., 2020; Bourven et al., 2015; Felz et al., 2020; van Vliet et al., 2020; Xue et al., 2019). This suggests that the production of sulfated polymers could occur within the biofilm, and is potentially linked with important functions in the biofilm. However, previous studies targeted biofilms in natural environments or microcosms where organic substrates containing sulfated polysaccharides were present and/or supplied. Thus, the question arises as to the metabolic potential of prokaryotic cells to synthesize sulfated polymers.

Our research aims to investigate if sGAG-like polymers can be microbially produced. For this we studied a microbial biofilm growing in the Puturosu Sulfur Cave (Romania), an extreme environment where the possibility of obtaining sGAG-like polymers from the environment is very unlikely. In the Puturosu Sulfur Cave, biofilms grow on the cave walls at the level of the interface between volcanic gas (CO_2, CH_4, and H_2S) and overlying atmospheric air. Stable isotope measurements showed that the organic carbon is autotrophic in nature (Sarbu et al., 2018). Deposits of elemental sulfur are present on the walls below this gas-gas interface. The measured pH on the cave wall range between 0.5 and 1, indicating possible sulfide oxidation to sulfate. No direct water source was found in the cave, implying that water is only available through condensation inside the cave (Sarbu et al., 2018). This avoids external input of organic sulfated polymers from the environment to the biofilm. If the sGAG-like polymers are detected in this biofilm, this hence supports the hypothesis that sulfated polymers are produced *in situ* by the microorganisms. This hypothesis was tested by measuring the presence of sGAG in the EPS of biofilm growing at the gas-gas interface. EPS was extracted from the biofilm and highly negatively charged polymers were isolated and detected using a specific

cationic dye for sGAG. Further validation of sGAG presence was performed with enzymatic treatment with chondroitinase ABC and FT-IR measurements of the biofilm, EPS, pellet and dye precipitated fraction. Lastly, the possibility of the microorganisms in the biofilm to produce sulfated polymers was explored by performing genome database searches for sulfotransferase, which is an important enzyme needed for sGAG-like polymer production. The detection of sGAG-like polymers in this extreme environment provides support that sulfated polymers can be produced exclusively by microorganisms and could be an important step in understanding the role of specific EPS compounds in the biofilm.

2. Materials and Methods

2.1. Biofilm sampling

Biofilm was sampled from Sulfur Cave located on Puturosu Mountain ("Stinky Mountain" in English translation), located within the Ciomadul volcanic edifice area in the East Carpathian Mountains of Romania. Sulfur Cave provides a unique extreme environment, as it displays microbial life within a gas chemocline (Figure. 1a). The cave penetrates approximately 14 meters into the volcanic bedrock and shows a gradual downward slope. The bottom section of the cave is filled by a continuous inflow of 2000 m^3/day with an anoxic gaseous mixture (CO_2, CH_4, and H_2S), while an ambient air layer (containing O_2) floats on top of the heavier gases (Sarbu et al., 2018). At the back end of the cave, the depth of the anoxic gas layer is the highest and reaches about 2 meters. Elemental sulfur (S°) and sulfuric acid (H_2SO_4) cover the cave walls at and below this gas chemocline. The pH drops to very low values (<1) near the chemocline, suggesting oxidation of H_2S. At and above the chemocline, microbial biofilms are present on the cave wall, which can be described as a white/brown colored slimy layer ranging in thickness of a few mm (Sarbu et al., 2018).

Eight biofilm samples were collected from the cave wall at the back end of the cave (Figure 1a). Samples were collected by scraping the biofilm from the cave wall into sterile eppendorf tubes (2 ml), which were closed inside the cave. All samples were frozen (-20°C) until further processing. One intact biofilm sample was imaged with JEOL 5600 scanning electron microscope (SEM) operated at 10 keV and 14500x magnification under low vacuum (Figure 1b).

2.2. Extraction of extracellular polymeric substances

In the lab, seven biofilm samples were lyophilized. Extraction of extracellular polymeric substances (EPS) was performed on the lyophilized samples following an alkaline extraction protocol (Pinel et al., 2020). In brief, 10 mg of lyophilized sample were added to 1 mL of 0.1 M NaOH solution and stirred vigorously for 30 minutes at 80 °C. Then, mixture was cooled to 4 °C and centrifuged at 4000 rpm for 30 minutes. The supernatant was dialyzed with a 3.5 kDa cut-off dialysis bag against demi-water overnight at room temperature and, subsequently, lyophilized. The EPS extraction yield was determined by dividing the weight of lyophilized supernatant by the weight of the lyophilized biofilm sample.

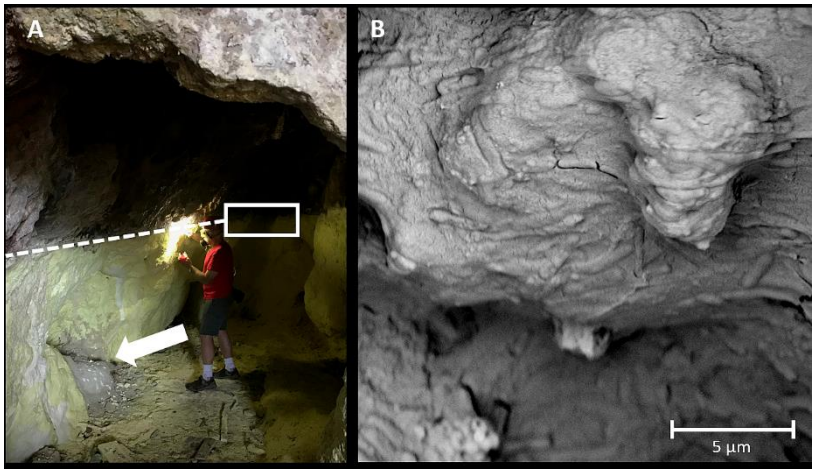


Figure 1. A) Sulfur Cave with white rectangle showing sampling location. The grey zone/white arrow at the bottom of the wall is the point where the volcanic gas emanates. Yellow deposits are sulfur deposits and show the height of the gas/gas interface (indicated by dashed white line). B) Scanning electron microscope (SEM) image of collected biofilm where numerous rod shaped cells embedded in the EPS are visible.

2.3. Extracellular polymeric substances characterization

2.3.1. Protein and carbohydrate analysis

The protein content of two EPS samples was measured in triplicate with a commercially available protein quantification kit (BC assay protein quantification kit, Interchim, Montluçon FRANCE) with bovine serum albumin (BSA) as the standard. Samples were dissolved in 0.01 M NaOH to a 2 mg/mL concentration and protein quantified according to the manufacturer's instructions. Absorbance was measured at 562 nm using a multimode plate reader (TECAN Infinite M200 PRO). The sugar content of the same samples was measured in triplicate with a phenol-sulfuric acid assay (Dubois et al. 1956), by measuring the absorbance at 482 nm with a VIS-spectrophotometer (HACH DR3900). A mixture of glucose, xylose, fucose, rhamnose, galactose, glucose, xylose, mannose and ribose was used as standard to ensure accurate quantification of sugar concentrations (see detailed explanation in Felz et al. 2019).

2.3.2. FT-IR analysis

FT-IR spectra were collected for 5 samples: one intact biofilm, two recovered EPS samples, one pellet (after EPS extraction) and one DMMB positive fraction (see section 2.3.3 below). Spectra were recorded with a Perkin-Elmer spectrum 100 FTIR spectrometer (Perkin-Elmer, Shelton, USA), over the wavenumber range from 4000 cm^{-1} to 650 cm^{-1} with 32 accumulations and 4 cm^{-1} resolution. Spectral data processing was performed in MATLAB and consisted of baseline correction (linear baseline correction) and feature scaling of the individual spectra.

2.3.3. sGAG-like polymers quantification

The total carbohydrate assay (e.g. phenol-sulfuric acid assay in section 2.3.1) does not extract a number of carbohydrates, including amino sugars, sialic acids and sulfated glycosaminoglycans (sGAG), as those carbohydrates do not form a furfural structure when they react with concentrated sulfuric acids (Felz et al., 2019; de Graaf et al., 2019). Dye-spectrometry with DMMB a cationic dye is a conventional method for measuring polyanionic carbohydrates like sGAG (Kubaski et al., 2017). It is a simple and fast method for the detecting the presence of sGAG, but no additional structural information is available. This is why additional analysis were added to validate the presence of sGAG-like polymers in the acidophilic biofilm. The presence of polyanionic carbohydrates such as sulfated glycosaminoglycans was evaluated in the intact biofilm (one sample) and in the extracted EPS (two samples) by conventional DMMB dye-spectrometry with the Blyscan™ glycosaminoglycan assay according to the manufacturer's instructions (Biocolor, UK). In brief, 20-50 mg of samples were treated with papain protein digestion solution (Sigma Aldrich) overnight. The supernatant was collected after centrifuging at 13000 rpm for 10 min. Supernatant (50 µL) was mixed with 1 mL of DMMB-dye reagent, in triplicate per sample. The DMMB dye is a cationic dye with a pH of 1.7 to selectively bind sulfated groups (Boleij et al., 2020). It forms precipitates with sulfated polysaccharides (e.g. sGAG). After resolubilization, the DMMB bound sGAG was quantified by measuring its absorbance at 656 nm with a multimode plate reader (TECAN Infinite M200 PRO). Chondroitin-6-sulfate is used as the standard. For the FT-IR measurement, the sample which had the highest concentration of sGAG-like polymers was selected. 200 µL of the remaining supernatant was mixed with 4 mL of DMMB-dye reagent. The sGAG-DMMB precipitates were lyophilized for FT-IR analysis.

2.3.4. DNase and Chondroitinase ABC digestion of enriched sulfated glycosaminoglycan-like polymer

The Blyscan™ glycosaminoglycan assay is designed to be selective for sGAGs like chondroitin-6-sulfate. Other polyanions can interfere and cause a false positive signal. Therefore the effect of enzymatic digestion was used to further investigate any structural similarity between DMMB positive fraction of the EPS and chondroitin-6-sulfate, EPS samples were digested with chondroitinase ABC (ChABCase from *Proteus vulgaris* (Sigma-Aldrich, Zwijndrecht, Netherlands)). ChABCase catalyzes the cleavage of chondroitin-4 sulfate, chondroitin-6 sulfate, dermatan sulfate and to a lesser extent hyaluronic acid. The sGAG concentration was measured after enzymatic treatment to determine the effectiveness of the enzymatic digestion, and thus the similarity of the sulfated polymer to the chondroitin-6-sulfate standard.

The presence of DNA may cause interference to sGAG quantification as it is also a polyanionic molecule. Thus, a DNase I treatment was performed to evaluate the interference of DNA. One papain digested sample was divided into three fractions, with one fraction as control and the other two fractions treated by chondroitinase ABC and DNase I, respectively, according to Zheng and Levenston (2015).

Chondroitinase ABC was prepared at 20U/mL in 0.01% BSA, 50 mM Tris and 60 mM sodium acetate at pH 8 (Felz et al., 2020). DNase I was prepared at 400U/mL working solution in 10 mM Tris-HCl, pH 7.5, 10 mM $CaCl_2$, 10 mM $MgCl_2$ and 50% (v/v) glycerol (Zheng and Levenston, 2015). For the enzymatic treatment, 90 µL of papain digested samples were digested with 5 µL of enzyme working solution. After enzymatic treatment, sGAG concentration was measured with Blyscan™.

2.4. BLAST analysis of enzymes related to sulfated glycosaminoglycans synthesis

The Basic Local Alignment Search Tool (BLAST) was used, with blastp from the NCBI website to find homologs of sulfotransferases, an enzyme known to be used by eukaryotes to synthesize sGAGs. To this end, two reference protein sequences from *Pseudomonas ogarae* (NCBI:txid1114970) and from *Sinorhizobium fredii* (NCBI:txid380) were matched against a protein database that includes the NCBI sequences from four major phylotypes (*Mycobacterium* (NCBI:txid1763), *Acidomyces* (NBCI:txid245561), *Acidithiobacillus* (NCBI:txid119977), *Ferroplasmaceae* (NCBI:txid90142)) that were previously shown to be present in the acidophilic biofilm of Sulfur Cave (Sarbu et al., 2018). The reference protein sequences were chondroitin 4-O sulfotransferase from *Pseudomonas ogarae* (accession number: WP_014336261.1) and heparan sulfate glucosamine 3-O sulfotransferase from *Sinorhizobium fredii* (accession number: WP_015887312.1)(Boleij et al., 2020). The matched sequences were considered significantly similar when the *E*-value was below 1E-8.

3. Results

3.1. EPS, carbohydrate and protein content of Sulfur Cave biofilms

During EPS extraction of the biofilm samples, the biological fraction was solubilized and separated from mineral components (pellet). Through dialysis, polymers larger than 3.5 kDa were separated from smaller molecules. The average EPS extraction yield was 5.0 ± 2.7% of the lyophilized biofilm weight. The insoluble fraction amounted to an average of 62.7 ± 1.0% of the lyophilized biofilm weight (supplemental table 1). The remaining 32.3% were dissolved and smaller than 3.5 kDa, and thus passed through the dialysis bag. The total protein and carbohydrate content in the extracted EPS were on average 32.3 ± 0.5% (BSA equivalents) and 26.3 ± 6.5% (carbohydrate standard mix equivalents), respectively. Underestimation of the polysaccharides in the extracted EPS can occur due to inability to measure amino sugars with the phenol-sulfuric acid method, and this underestimation could explain part of the unquantified fraction of EPS (Felz et al., 2019). Other EPS components can be e.g. lipids and humic substances (Flemming and Wingender, 2010).

3.2. FT-IR analysis of Pellet, Biofilm and EPS fractions

Table 1. Overview of absorbance bands with their respective spectral assignments.

band (cm^{-1})	assignment	reference
720	Si-Si(Al) and Si-Si stretch	(Chen et al., 2015; Theodosoglou et al., 2010).
770	Si-Si(Al) and Si-Si stretch	(Chen et al., 2015; Theodosoglou et al., 2010).
1005	Si(Al) stretch	(Chen et al., 2015; Theodosoglou et al., 2010).
1040-1160	C-O stretch (proteins and carbohydrates)	(Talari et al., 2017)
1222-1265	Phosphate asymmetric stretch	(Talari et al., 2017)
1230-1260	S=O asymmetric stretch	(Cabassi et al., 1978; Devlin et al., 2019)
1540-1570	Amide II (protein)	(Talari et al., 2017)
1620-1680	Amide I (protein)	(Talari et al., 2017)
2921-2923	CH_2 symmetric stretch	(Talari et al., 2017)
3200-3500	OH-stretch, Amide A	(Talari et al., 2017)

FT-IR spectroscopy was used to identify and compare main components in the insoluble pellet, intact biofilm and EPS fractions (Figure 2). Organic and inorganic functional groups could be assigned from the measured spectra (Table 1). In the insoluble pellet, a low absorbance (0.02 a.u.) was measured in amide I bands, which can be used as an indicator for organic components. Inorganic fractions were more prevalent as the normalized absorbance of inorganic functional groups reached 1.0 (a.u.), thus implying that the insoluble pellet contains mainly minerals.

The spectrum of the biofilm displayed high normalized absorbance both at the regions indicating mineral (0.75 a.u.), and polysaccharide components (1.0 a.u.), however overlap of these bands interfere with functional group assignment. Alternatively, higher absorbance in the amide I protein band and OH polymeric stretch of the biofilm spectrum (0.14 and 0.14 a.u., respectively), suggested more organic components compared to the insoluble pellet (0.02 and 0.02 a.u., respectively). The EPS spectrum showed that the extraction is successful in extracting organic polymeric components from the biofilm. The absorbance of protein bands and OH-polymeric stretch increased from 0.14 to 0.56 and 0.14 to 0.38 a.u., respectively. Additionally, the polysaccharide absorbance peak was sharper and a shoulder band indicating possible sulfate presence became apparent.

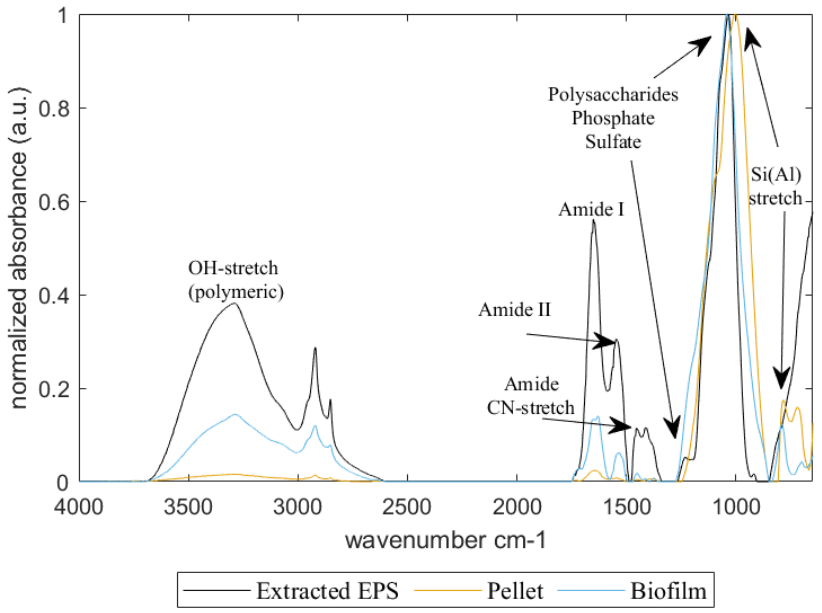


Figure 2. Normalized FT-IR spectra of insoluble pellet, intact biofilm and extracted EPS. Spectra were recorded from 4000 to 650 cm^{-1} and measured absorbances normalized by dividing by the maximum absorbance (shown in the figure as arbitrary unit a.u).

3.3. Detection of sulfated-glycosaminoglycans-like polymers

The detection and quantification of the sulfated polymers was performed using the Blyscan™ glycosaminoglycan assay. Considering the extreme low pH in the cave, the appearance of strong polyanionic carbohydrates, with strong acidic groups, can be expected in the biofilm. Therefore, the occurrence of polyanionic carbohydrates such as sulfated glycosaminoglycans (sGAG) in the biofilm was evaluated. The sGAG concentration was measured in triplicate on one biofilm sample that did not undergo EPS extraction. On average, 5.45 ± 0.02 μg/mg of the lyophilized biofilm material was composed of sGAG-like polymers (chondroitin sulfate equivalent). In the extracted EPS, the content of sGAG-like polymers 21.5 and 56.2 μg/mg EPS for the two samples analyzed.

Since the quantification assay is based on precipitation of the DMMB–sulfated glycosaminoglycan complex, it also served as a method for sGAG-like polymer isolation. In theory, the isolated DMMB positive fraction is enriched with sGAG-like polymer. To validate this, the FT-IR spectra of the recovered EPS and the DMMB positive fraction were compared (Figure 3). Both spectra contained bands around 1023 and 1047 cm^{-1}, implying C–O stretch in carbohydrates. The intensity of absorbance bands of OH-polymeric stretch and amide I and II bands, is decreased

in the spectrum of DMMB positive fraction in comparison to that of extracted EPS, which is due to the removal of proteins during the pretreatment of DMMB assay. Interestingly, bands at around 1212 and 1167 cm^{-1} in the spectrum of DMMB positive fraction became apparent, which are indicators for sulfate half-esters and polysaccharide, respectively (Bombalska et al., 2011).

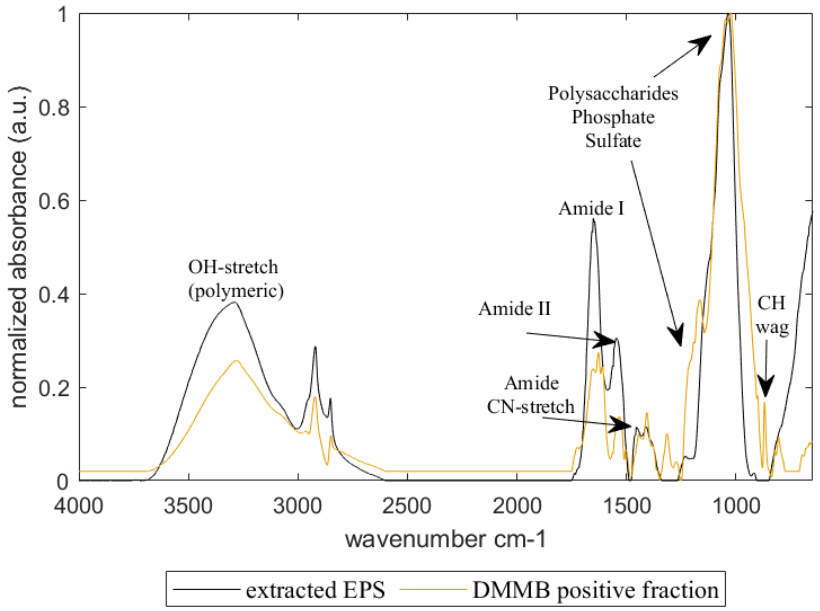


Figure 3. Normalized FT-IR spectra of extracted EPS and complexed sulfated polymers. Spectra were recorded from 4000 to 650 cm^{-1} and normalized by rescaling to maximum absorbance, absorbance is in arbitrary units (a.u).

3.4. Enzymatic digestion of enriched sulfated glycosaminoglycans-like polymer

To further confirm the structural similarity between DMMB positive fraction and chondroitin sulfate, Chondroitinase ABC digestion was conducted, which resulted in a 26.3 ± 7.9% signal decrease after measurement with Blyscan™ sGAG kit. Thus chondroitinase could effectively digest part of the DMMB positive fraction. This implies that there is a similar structure with chondroitin-sulfate. If there is DNA in the sample, it can potentially bind to DMMB, resulting in an overestimation of sGAG-like polymers (Zheng and Levenston, 2015). However, DNase treatment of the samples resulted in a 3.3 ± 1.9% signal decrease of the DMMB positive fraction only, meaning a minor interference of DNA in the estimation of the total sGAG-like polymers.

3.5. Mining prokaryotic protein databases for sulfotransferases

Sulfotransferases are enzymes that transfer sulfo- groups to polysaccharides (Kusche-Gullberg and Kjellén, 2003) and are an important factor in the biosynthesis of sGAGs. Knowledge on the occurrence of sulfotransferases in prokaryotes is limited; here we mined protein databases of microorganisms belonging to the genera *Mycobacterium*, *Acidomyces* and *Acidithiobacillus* for sulfotransferases. These genera were predominantly found in samples for the Sulfur Cave analyzed previously by Sarbu et al. (2018). Two reference proteins were used for protein Blast (BlastP): chondroitin 4-O sulfotransferase from *Pseudomonas fluorescens* and heparan sulfate glucosamine 3-O-sulfotransferase from *Sinorhizobium fredii* (Boleij et al., 2020). For chondroitin 4-O sulfotransferase, two matches were obtained, both occurring in the *Mycobacterium* genus (Table 2). For the second enzyme, heparan sulfate glucosamine 3-O-sulfotransferase, forty protein matches were found in the *Acidithiobacillus* and *Mycobacterium* genera (Table 3). These results suggest that at least two of the microbial genera reported to inhabit the Sulfur Cave have the potential to synthesize sulfated polysaccharides.

Table 2. List of significant BlastP search matches against chondroitin 4-O sulfotransferase (accession number: WP_014336261.1)

Accession number	Name protein and species	Identity (%)	Bitscore	E-value
WP_207065469.1	sulfotransferase family 2 domain-containing protein [unclassified mycobacterium]	31.016	225	1.31e-15
WP_192722148.1	sulfotransferase family 2 domain-containing protein [Mycobacterium sp. OAS707]	31.056	236	2.68e-08

Table 3. List of significant BLAST search matches against heparan sulfate glucosamine 3-O sulfotransferase (accession number: WP_015887312.1)

Accession number	**Name protein and [species]**	**Identity (%)**	**Bitscore**	**E-value**
WP_142086737.1	sulfotransferase [Acidithiobacillus thiooxidans]	28.767	291	4.21e-20

WP_163098766.1	sulfotransferase [Acidithiobacillus ferrianus]	32.161	216	1.62e-19
WP_163055154.1	sulfotransferase [Acidithiobacillus ferrooxidans]	29.825	291	2.99e-19
WP_192722155.1	sulfotransferase domain-containing protein [Mycobacterium sp. OAS707]	31.944	290	3.18e-19
WP_126605524.1	MULTISPECIES: sulfotransferase [Acidithiobacillus]	28.772	291	7.68e-19
WP_012536779.1	MULTISPECIES: sulfotransferase [Acidithiobacillus]	28.772	291	8.24e-19
WP_113525668.1	sulfotransferase [Acidithiobacillus ferridurans]	28.772	291	1.06e-18
WP_014029380.1	sulfotransferase [Acidithiobacillus ferrivorans]	29.123	291	1.25e-18
WP_035190410.1	sulfotransferase [Acidithiobacillus ferrivorans]	29.123	291	1.43e-18
WP_193650373.1	MULTISPECIES: sulfotransferase [unclassified Acidithiobacillus]	28.814	303	1.58e-18
WP_101537278.1	sulfotransferase [Acidithiobacillus sp. SH]	28.239	301	1.79e-18
WP_010638122.1	sulfotransferase [Acidithiobacillus thiooxidans]	31.088	200	3.20e-18
WP_065973527.1	sulfotransferase [Acidithiobacillus thiooxidans]	31.088	200	4.11e-18

WP_031568475.1	MULTISPECIES: sulfotransferase [Acidithiobacillus]	31.088	200	4.40e-18
QFX96614.1	sulfotransferase [Acidithiobacillus thiooxidans ATCC 19377]	31.088	202	5.20e-18
WP_206385958.1	sulfotransferase [Acidithiobacillus sp. MC6.1]	28.772	291	6.58e-18
WP_176562256.1	sulfotransferase domain-containing protein [Mycobacterium palauense]	34.254	187	6.58e-18
WP_163098168.1	sulfotransferase [Acidithiobacillus ferrianus]	28.772	291	8.35e-18
TQN51840.1	hypothetical protein DLNHIDIE_01720 [Acidithiobacillus thiooxidans ATCC 19377]	28.571	303	8.50e-18
WP_193870059.1	sulfotransferase [Acidithiobacillus sulfuriphilus]	30.964	204	9.58e-18
WP_064219950.1		28.070	291	1.52e-17
WP_140390866.1	MULTISPECIES: sulfotransferase [Acidithiobacillus]	28.231	303	1.61e-17
WP_024895070.1	sulfotransferase [Acidithiobacillus thiooxidans]	28.475	303	1.91e-17
WP_024892770.1	sulfotransferase [Acidithiobacillus thiooxidans]	30.570	200	2.09e-17
WP_193650115.1	MULTISPECIES: sulfotransferase [unclassified Acidithiobacillus]	30.570	200	4.60e-17

WP_075323453.1	sulfotransferase [Acidithiobacillus albertensis]	27.891	303	5.26e-17
WP_010639455.1	sulfotransferase [Acidithiobacillus thiooxidans]	30.808	216	1.42e-16
WP_193640087.1	sulfotransferase [Acidithiobacillus sp. HP-11]	29.851	215	3.06e-16
HGE68600.1	TPA: sulfotransferase [Acidithiobacillus sp.]	32.768	182	6.24e-16
WP_142089850.1	sulfotransferase [Acidithiobacillus thiooxidans]	29.353	215	6.75e-16
WP_142087470.1	sulfotransferase [Acidithiobacillus thiooxidans]	29.353	215	7.30e-16
WP_193651055.1	MULTISPECIES: sulfotransferase [unclassified Acidithiobacillus]	29.353	215	7.45e-16
WP_004871655.1	sulfotransferase [Acidithiobacillus caldus]	33.696	201	1.50e-15
WP_014002827.1	sulfotransferase [Acidithiobacillus caldus]	32.609	201	1.93e-15
WP_070114154.1	sulfotransferase [Acidithiobacillus caldus]	32.609	201	2.19e-15
WP_198926362.1	sulfotransferase [Acidithiobacillus thiooxidans]	27.972	278	1.66e-14
MBO0863535.1	sulfotransferase domain-containing protein [Mycobacterium sp.]	28.936	265	2.90e-14
WP_066916256.1	sulfotransferase [Mycobacterium interjectum]	30.769	186	1.69e-10

SPM30343.1	putative sulfotransferase, partial [Mycobacterium terramassiliense]	30.220	186	1.94e-10
WP_077101318.1	sulfotransferase [Mycobacterium terramassiliense]	30.220	186	2.07e-10

4. Discussion

Currently, the microbial production of glycosaminoglycan-like polymers is mostly studied in pathogenic bacteria, e.g. *Pseudomonas aeruginosa* and model organisms like *Escherichia coli*. These bacteria produce non-sulfated glycosaminoglycans, or closely related molecules, which mimic the backbone of sulfated glycosaminoglycans produced by the host, supposedly as a camouflage to avoid detection by the host immune system (Badri et al., 2021). Recently, sGAG-like polymers have been reported in prokaryotic ecosystems, such as marine microbial mats (Decho and Gutierrez, 2017) and granular biofilms (including aerobic, anaerobic and anammox granular sludge) collected from municipal wastewater treatment plants (Felz et al., 2020; Bourven et al., 2015, Boleij et al., 2020). This study targets another type of biofilm ecosystems, i.e., extremely acidophilic biofilms.

Our results show the presence of sGAG-like polymers in the acidophilic biofilm growing in Sulfur Cave. Firstly, the results from the DMMB assay indicated the presence of anionic polymers. Since the Blyscan kit is designed for quantification of sGAG, it is selective for strong acidic polymers. Secondly, the enriched polymer was sensitive to the digestion with chondroitinase ABC, implying a similar but not necessary identical structure as chondroitin sulfate is present in the biofilm. It is known that the cleavage sites of chondroitinase are: (1→3) glycosidic linkage between β-D-glucuronic acid and N-Acetyl-β-D-galactosamine-4-sulfate; (1→3) glycosidic linkage between β-D-glucuronic acid and N-Acetyl-β-D-galactosamine-6-sulfate; and (1→3) and (1→4) glycosidic linkages between β-D-glucuronic acid and N-Acetyl-β-D-glucosamine (Sigma-Aldrich, 2007; Felz et al., 2020). So the molecular structure of the enriched polymer should have at least one of those linkages. Thirdly, the FT-IR spectrum of the DMMB positive fraction is dominated by absorbance bands in the polysaccharide region (1200-900 cm^{-1}). Two bands in particular were more pronounced in the spectrum of the DMMB positive fraction relative to that of EPS. These two bands suggest the possible presence of organosulfate, and/or the C–O–C stretch of uronic acids, respectively (Talari et al., 2016; Devlin et al., 2019). Therefore, the DMMB positive fraction is considered enriched with sGAG-like polymer. Lastly, when we screened the NCBI database, chondroitin 4-O sulfotransferase showed two hits in the *Mycobacterium* genus. As this enzyme determines the biosynthesis of chondroitin sulfate, this suggests the possibility of chondroitin sulfate (probably chondroitin 4-O sulfate) synthesis in the acidophilic biofilm. Additionally, the large number of microorganisms encoding heparan sulfate

glucosamine 3-O-sulfotransferase implies that there might be sGAGs other than chondroitin sulfate produced in the biofilm. Unlike the granular sludge collected from wastewater treatment plants, the biofilms growing in Sulfur Cave are not exposed to input of other organic substances. The detection of sGAGs-like polymers in this biofilm therefore provide strong support for the hypothesis that sGAGs-like polymers can be directly produced by prokaryotes.

Currently, the EPS sulfation process in microorganisms is still unclear. However, in a mammalian cell, the sulfation process of sGAG begins with uptake of inorganic sulfate from the extracellular milieu (de Costa et al., 2017). Thus to produce sGAGs, sulfate compounds should be readily available in the environment surrounding the biofilm. In this regard, sulfated extracellular carbohydrate-rich polymers are found in various marine biofilms. In fact, highly sulfated exopolysaccharides containing up to 27% (wt/wt) in sulfate were isolated from marine microbial mats (Moppert et al., 2009). The availability of sulfate might explain the prevalence of sulfated polymers in marine biofilms; sulfate is the second most abundant anion in the surrounding seawater, with an average concentration of 28 mM (Canfield and Farquhar, 2009). In biofilms used for treating wastewater (including activated sludge and granular sludge), the content of sulfate polysaccharides-containing EPS varies. For example, 418 mg sulfated EPS/g biomass was found in the sludge collected from full-scale wastewater treatment plant in Hong Kong, which contains 5.2 mM sulfate in the influent. Oppositely, no sulfated EPS was detected in the sludge collected from a wastewater treatment plant with no sulfate and/or sulfide in the influent (Xue et al., 2019). A lower content of sulfated EPS (24-31mg sGAG/g biomass) (Felz et al., 2020; Boleij et al., 2020) was further observed in biofilms used for treating non-saline domestic wastewater that had a lower sulfate concentration (0.28-0.62 mM ; van den Brand et al., 2015). Sulfur Cave is a high sulfate environment due to the oxidation of hydrogen sulfide to sulfuric acid (Sarbu et al., 2018). Assuming sulfuric acid is the only acid present, the maximum sulfate concentration in the cave can be estimated from the pH values ~0.5-1 as high as 100 mM. These high concentrations of sulfate may favor the biosynthesis of sGAGs. The Sulfur Cave biofilm however, contained sGAG-like polymers at 11 mg/g organic dry weight (estimated from biomass minus the pellet mineral fraction), which is rather low compared to the wastewater and marine microbial mat data previously published. The difference can be explained by a lower actual concentration of sulfate than estimated. Alternatively, in the wastewater and marine biofilms competition for sulfate is present in the form of sulfate reducing bacteria; which couple sulfate reduction to organic carbon oxidation. In these biofilms, charged modification, by e.g. sialic acids, of extracellular polymers provide protection from degradation (Traving and Schauer, 1998). Less competition of this kind is present in the Sulfur Cave due to a lower microbial diversity in this extreme environment (Sarbu et al., 2018). How microbial diversity and sulfate presence are connected to sulfated polymer production is yet unknown and could provide interesting insight in microbial interactions inside biofilms.

With regard to the function of sGAGs in the Sulfur Cave biofilm, one has to consider the low pH value (<1 pH) of the system. Under such extreme conditions, weakly acidic groups (carboxylic acids) become protonated and lose their charge

(Boleij et al., 2020; Felz et al., 2020), while highly acidic polymers, like sGAGs, may maintain a negatively charged EPS (Thornton et al., 2007; Gutierrez et al., 2009). A highly negative charged EPS contributes to the ability of the EPS to bind water or cations (Varki, A., 2017; Gagliano et al., 2018) and form a hydrogel. This hydrogel matrix not only helps microorganisms to compete and survive under extreme environmental conditions, but also maintain the stability of the biofilm. Protein to carbohydrate ratio from this unique biofilm (1.28) was lower than observed in granular sludge (2.61)(Felz et al., 2019). This difference might be due to differences in hydrophobicity. Granular sludge is typically hydrophobic, which is the result of a high protein content (Gao et al., 2011; Santschi et al., 2020). Inside Sulfur Cave, water is only provided by condensation, thus it would be advantageous for the microbes living in this environment to evolve a mechanism to retain water. Remarkably, the function of the microbial chondroitin sulfate in this acidophilic biofilm is likely similar as in animal cartilage where chondroitin sulfate is responsible for the water retention (Jerosch, 2011). In addition, chondroitin sulfate (sGAG) is stable at pH 1, only when it is heated at this pH at high temperature (>30°C) for more than 960 hours, a slight hydrolysis starts (Volpi et al., 1999). Thus the sGAG-like polymers found in the biofilm of Sulfur Cave would be stable at the low pH and may serve as a hydration mechanism. However, the exact function of sGAGs in biofilms requires targeted experiments and the extreme environment of Sulfur Cave offers a unique opportunity for in depth research on the topic.

5. Conclusion

This research set out to investigate the possible presence of sGAGs in a unique microbial biofilm isolated from any external organic carbon source. The acidophilic biofilm growing at the non-aquatic gas-gas interface of Sulfur Cave showed a clear presence of sulfated glycosaminoglycan-like polymers. Furthermore, the microbial taxa (*Mycobacterium* and *Acidithiobacillus*) present in the biofilm appear to have the metabolic potential to synthesize these compounds (chondroitin and heparan sulfate). The possibility of *de novo* microbial production of such polymers provides support for the widespread prevalence of sulfated glycosaminoglycan-like polymers in biofilms. The negatively charged nature of these polymers in fact suggests a potential functional role in water retention for the biofilm in Sulfur Cave, but more targeted research needs to be conducted to elucidate its true function.

Acknowledgements

The authors thank Jure Zlopasa for fruitful discussions and insight on negative charge compounds in biofilms. Additionally the authors thank the members of the Vinca Minor association for their efforts for maintaining and protecting the mofettes on Puturosu Mountain, and the GESS team for providing the protective equipment used during field work in the extreme conditions encountered at sites such as Sulfur Cave. This research was funded by the SIAM Gravitation Grant 024.002.002, The Netherlands Organization for Scientific Research. DVC and FJRM was financially supported by the Research Foundation Flanders via FWO grant G031416N, and the Netherlands Organization for Scientific Research (VICI grant 016.VICI.170.072).

References

Badri, A., Williams, A., Awofiranye, A., Datta, P., Xia, K., He, W., ... & Koffas, M. A. (2021). Complete biosynthesis of a sulfated chondroitin in Escherichia coli. *Nature communications*, *12*(1), 1-10.

Boleij, M., Kleikamp, H., Pabst, M., Neu, T. R., Van Loosdrecht, M. C., & Lin, Y. (2020). Decorating the Anammox House: Sialic Acids and Sulfated Glycosaminoglycans in the Extracellular Polymeric Substances of Anammox Granular Sludge. *Environmental science & technology*, *54*(8), 5218-5226.

Bombalska, A., Mularczyk-Oliwa, M., Kwaśny, M., Włodarski, M., Kaliszewski, M., Kopczyński, K., ... & Trafny, E. A. (2011). Classification of the biological material with use of FTIR spectroscopy and statistical analysis. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *78*(4), 1221-1226.

Bourven I, Bachellerie G, Costa G, Guibaud G. Evidence of glycoproteins and sulphated proteoglycan-like presence in extracellular polymeric substance from anaerobic granular sludge. Environ Technol. 2015;36(19):2428-35. doi: 10.1080/09593330.2015.1034186. Epub 2015 Apr 24. PMID: 25812669.

Cabassi, F., Casu, B., & Perlin, A. S. (1978). Infrared absorption and Raman scattering of sulfate groups of heparin and related glycosaminoglycans in aqueous solution. *Carbohydrate Research*, *63*, 1-11.

Canfield, D. E., & Farquhar, J. (2009). Animal evolution, bioturbation, and the sulfate concentration of the oceans. *Proceedings of the National Academy of Sciences*, *106*(20), 8123-8127.

Chen, Y., Zou, C., Mastalerz, M., Hu, S., Gasaway, C., & Tao, X. (2015). Applications of micro-fourier transform infrared spectroscopy (FTIR) in the geological sciences—a review. *International journal of molecular sciences*, *16*(12), 30223-30250.

de Graaff, D. R., Felz, S., Neu, T. R., Pronk, M., van Loosdrecht, M. C., & Lin, Y. (2019). Sialic acids in the extracellular polymeric substances of seawater-adapted aerobic granular sludge. *Water research*, *155*, 343-351.

Decho, A. W., & Gutierrez, T. (2017). Microbial extracellular polymeric substances (EPSs) in ocean systems. *Frontiers in microbiology*, *8*, 922.

Devlin, A., Mauri, L., Guerrini, M., Yates, E. A., & Skidmore, M. A. (2019). The use of ATR-FTIR spectroscopy to characterise crude heparin samples by composition and structural features. *bioRxiv*, 744532.

Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. T., & Smith, F. (1956). Colorimetric method for determination of sugars and related substances. *Analytical chemistry*, *28*(3), 350-356.

Felz, S., Neu, T. R., van Loosdrecht, M. C., & Lin, Y. (2020). Aerobic

granular sludge contains Hyaluronic acid-like and sulfated glycosaminoglycans-like polymers. *Water research*, *169*, 115291.

Felz, S., Vermeulen, P., van Loosdrecht, M. C., & Lin, Y. M. (2019). Chemical characterization methods for the analysis of structural extracellular polymeric substances (EPS). *Water Research*, *157*, 201-208.

Flemming, H. C., & Wingender, J. (2010). The biofilm matrix. *Nature reviews microbiology*, *8*(9), 623-633.

Gao, D., Liu, L., Liang, H., & Wu, W. M. (2011). Aerobic granular sludge: characterization, mechanism of granulation and application to wastewater treatment. *Critical reviews in biotechnology*, *31*(2), 137-152.

Gutierrez, T., Morris, G., & Green, D. H. (2009). Yield and physicochemical properties of EPS from Halomonas sp. strain TG39 identifies a role for protein and anionic residues (sulfate and phosphate) in emulsification of n-hexadecane. *Biotechnology and bioengineering*, *103*(1), 207-216.

Jerosch, J. (2011). Effects of glucosamine and chondroitin sulfate on cartilage metabolism in OA: outlook on other nutrient partners especially omega-3 fatty acids. *International journal of rheumatology*, *2011*.

Köwitsch, A., Zhou, G., & Groth, T. (2018). Medical application of glycosaminoglycans: a review. *Journal of tissue engineering and regenerative medicine*, *12*(1), e23-e41.

Kubaski, F., Osago, H., Mason, R. W., Yamaguchi, S., Kobayashi, H., Tsuchiya, M., ... & Tomatsu, S. (2017). Glycosaminoglycans detection methods: Applications of mass spectrometry. *Molecular genetics and metabolism*, *120*(1-2), 67-77.

Kusche-Gullberg, M., & Kjellén, L. (2003). Sulfotransferases in glycosaminoglycan biosynthesis. *Current opinion in structural biology*, *13*(5), 605-611.
Moppert, X., Le Costaouec, T., Raguenes, G., Courtois, A., Simon-Colin, C., Crassous, P., ... & Guezennec, J. (2009). Investigations into the uptake of copper, iron and selenium by a highly sulphated bacterial exopolysaccharide isolated from microbial mats. *Journal of Industrial Microbiology and Biotechnology*, *36*(4), 599-604.

Pinel, I. S., Kleikamp, H. B., Pabst, M., Vrouwenvelder, J. S., van Loosdrecht, M., & Lin, Y. (2020). Sialic acids: An important family of carbohydrates overlooked in environmental biofilms. *Applied Sciences*, *10*(21), 7694.

Santschi, P. H., Xu, C., Schwehr, K. A., Lin, P., Sun, L., Chin, W. C., ... & Quigg, A. (2020). Can the protein/carbohydrate (P/C) ratio of exopolymeric substances (EPS) be used as a proxy for their 'stickiness' and aggregation propensity?. *Marine Chemistry*, *218*, 103734.

Sarbu, S. M., Aerts, J. W., Flot, J. F., Van Spanning, R. J., Baciu, C., Ionescu, A., ... & Hegyeli, B. (2018). Sulfur Cave (Romania), an extreme environment with microbial mats in a CO 2-H 2 S/O 2 gas chemocline dominated by mycobacteria. *International Journal of Speleology*, *47*(2), 7.

Schiraldi, C., Cimini, D., & De Rosa, M. (2010). Production of chondroitin sulfate and chondroitin. *Applied microbiology and biotechnology*, *87*(4), 1209-1220.

Soares da Costa, D., Reis, R. L., & Pashkuleva, I. (2017). Sulfation of glycosaminoglycans and its implications in human health and disorders. *Annual review of biomedical engineering*, *19*, 1-26.

Talari, A. C. S., Martinez, M. A. G., Movasaghi, Z., Rehman, S., & Rehman, I. U. (2017). Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews*, *52*(5), 456-506.

Theodosoglou, E., Koroneos, A., Soldatos, T., Zorba, T., & Paraskevopoulos, K. M. (2010). Comparative Fourier transform infrared and X-ray powder diffraction analysis of naturally occurred k-feldspars. *Bulletin of the Geological Society of Greece*, *43*(5), 2752-2761.

Thornton, D. C., Fejes, E. M., DiMarco, S. F., & Clancy, K. M. (2007). Measurement of acid polysaccharides in marine and freshwater samples using alcian blue. *Limnology and Oceanography: Methods*, *5*(2), 73-87.

Traving, C., & Schauer, R. (1998). Structure, function and metabolism of sialic acids. *Cellular and Molecular Life Sciences CMLS*, *54*(12), 1330-1349.

van Vliet, D. M., Lin, Y., Bale, N. J., Koenen, M., Villanueva, L., Stams, A. J., & Sánchez-Andrea, I. (2020). Pontiella desulfatans gen. nov., sp. nov., and Pontiella sulfatireligans sp. nov., Two Marine Anaerobes of the Pontiellaceae fam. nov. Producing Sulfated Glycosaminoglycan-like Exopolymers. *Microorganisms*, *8*(6), 920.

Varki, A. (2017). Biological roles of glycans. *Glycobiology*, *27*(1), 3-49.

Volpi, N., Mucci, A., & Schenetti, L. (1999). Stability studies of chondroitin sulfate. *Carbohydrate research*, *315*(3-4), 345-349.

Xue, W., Zeng, Q., Lin, S., Zan, F., Hao, T., Lin, Y., ... & Chen, G. (2019). Recovery of high-value and scarce resources from biological wastewater treatment: Sulfated polysaccharides. *Water research*, *163*, 114889.

Zheng, C., & Levenston, M. E. (2015). Fact versus artifact: avoiding erroneous estimates of sulfated glycosaminoglycan content using the dimethylmethylene blue colorimetric assay for tissue-engineered constructs. *European cells & materials*, *29*, 224.

Zykwinska, A., Marchand, L., Bonnetot, S., Sinquin, C., Colliec-Jouault, S., & Delbarre-Ladrat, C. (2019). Deep-sea Hydrothermal Vent Bacteria as a Source of Glycosaminoglycan-Mimetic Exopolysaccharides. *Molecules*, *24*(9), 1703.

Supplemental information

Supplemental table 1. Overview of extraction yields on samples and protein and carbohydrate content.

Biofilm Sample	Dry weight (DW) (mg)	Pellet fraction (%, mg/mgDW)	EPS fraction (%, mg/mgDW)	Protein fraction ($mg_{\mathbf{protein}}/mg_{\mathbf{EPS}}$)	Carbohydrate fraction (%, $mg_{carbohydrate}/mg_{EPS}$)
S1	27.3	n.m.	3.3	29.7	22.1
S2	121	69.7	3.4	35.0	28.6
S3	160	69.1	6.2	n.m.	n.m.
S4	32.9	46.8	4.4	n.m.	n.m.
S5	83.1	62.1	10.5	n.m.	n.m.
S6	122.5	72.8	2.6	n.m.	n.m.
S7	94.2	55.8	4.6	n.m.	n.m.
Average		62.7 ± 9.9	5.0 ± 2.7		

n.m.: not measured

Chapter 3: Sialylation and sulfation of anionic glycoconjugates are common in the extracellular polymeric substances of both aerobic and anaerobic granular sludge

Abstract
Anaerobic and aerobic granular sludge processes are widely applied in wastewater treatment. In these systems, microorganisms grow in dense aggregates due to the production of extracellular polymeric substances (EPS). This study investigates the sialylation and sulfation of anionic glyconconjugates in anaerobic and aerobic granular sludges collected from full-scale wastewater treatment processes. Size exclusion chromatography revealed a wide molecular weight distribution (3.5 - >5,500 kDa) of the alkaline-extracted EPS. The high molecular weight fraction (>5,500 kDa), comprising 16.9 - 27.4% of the EPS, was dominant with glycoconjugates. Mass spectrometry analysis and quantification assays identified nonulosonic acids (NulOs, e.g. bacterial sialic acids) and sulfated groups contributing to the negative charge in all EPS fractions. NulOs were predominantly present in the high molecular weight fraction (47.2 - 84.3% of all detected NulOs), while sulfated glycoconjugates were distributed across the molecular weight fractions. Microorganisms, closely related to genera found in the granular sludge communities, contained genes responsible for NulOs and sulfate group synthesis or transfer. The similar distribution pattern of sialylation and sulfation of the anionic glycoconjugates in the EPS samples indicates that these two glycoconjugate modifications commonly occur in the EPS of aerobic and anaerobic granular sludges.

Keywords:

EPS, biofilm, size exclusion chromatography, mass spectrometry, nonulosonic acids, glycoconjugates

Synopsis:

This study provides insight into the distribution of potential valuable components in the extracellular matrix of granular sludge, thereby contributing to waste stream valorization.

Published as:

Chen, L. M., de Bruin, S., Pronk, M., Sousa, D. Z., van Loosdrecht, M. C., & Lin, Y. (2023). Sialylation and sulfation of anionic glycoconjugates are common in the extracellular polymeric substances of both aerobic and anaerobic granular sludges. Environmental Science & Technology, 57(35), 13217-13225.

1. Introduction

Microbial granulation is often desired in wastewater treatment processes, as the higher sedimentation velocity of granular sludge allows the ease of biomass separation from treated water (Pronk et al., 2015; Van Lier et al., 2008). Both aerobic and anaerobic microorganisms can granulate by immobilization in a matrix of self-excreted extracellular polymeric substances (EPS). In both processes, the technology inherently relies on the stability of the granules and the formation of EPS. The importance of negatively charged groups in EPS for granule stability and for the adsorption of charged substances has been highlighted (Costa et al., 2018). Some studies have focused on negatively charged polysaccharides, as for example bacterial alginate (Flemming & Wingender, 2010). However, negatively charged groups can also be found on the glycoconjugates linked to proteins and/ or lipids in EPS, e.g., sialic acids and sulfated groups.

Sialic acids are nine-carbon acidic monosaccharides that are mostly detected on the terminal of the glycoconjugate chain in the extracellular matrix of vertebrate cells or pathogenic bacteria (Varki et al., 2017). The most common sialic acids in animal tissue are N-acetyl neuraminic acid (NeuAc) and 2-keto-deoxynonulosonic acid (KDN), whereas pseudaminic acid (Pse) and its stereoisomer legionaminic acid (Leg) seem exclusive bacterial sialic acids (Morrison et al., 2014; Kleikamp et al., 2020). These are all monosaccharides belonging to a subset of the family of nonulosonic acids (NulOs). Most literature reports on sialic acids have focused on their role in evolution and disease in vertebrates, or the interaction between host cells and pathogenic bacteria (Varki et al., 2017). Only very recently, the presence of NulOs in several non-pathogenic microorganisms has been described. NeuAc was found in a diversity of environmental samples and associated to non-pathogenic microbial species (Kleikamp et al., 2020; de Graaff et al., 2019; Boleij et al., 2020). Leg/Pse were predominant in an enrichment of the phosphate-accumulating organism, "*Candidatus* Accumulibacter phosphatis" (Tomás-Martinez et al., 2021). In S-layer glycoprotein of the Archaea *Halorubrum sp* PV6, Leg was detected and speculated to be important for cell-cell recognition (Zaretsky et al., 2018). It is therefore suggested that glycoconjugates containing these monosaccharides (glycoconjugates with sialylation) may play a role in microbial aggregates, where microbe-microbe interactions occur.

Sulfated groups have been well-studied in mucin and the proteoglycan component in the extracellular matrix of animals, especially in sulfated glycosaminoglycans (GAGs). Sulfated GAGs are highly negatively charged, linear polysaccharide chains, covalently linked to the protein core (Varki et al., 2017). They are involved in distinct functions such as keeping the structural integrity of the extracellular matrix, wound repairing and cell differentiation in eukaryotes. Like sialic acids, sulfated GAGs have been believed to be produced mostly by pathogenic bacteria. However, recently, the sulfated GAGs have been found in the capsule surround the microorganisms and the EPS between the microcolonies in aerobic and anammox granules (Boleij et al., 2020; Felz et al., 2020). In anaerobic granular sludge, sulfated proteoglycan-like compounds have been reported (Bourven et al., 2015).

Despite being carbohydrates, sialic acids and sulfated glycoconjugates cannot be detected by frequently used carbohydrate assay, which contributed to a so far underestimation of their occurrence, chemical structures, and location in the EPS (de Graaff et al., 2019; Felz et al., 2020; Masuko et al., 2005; Parc et al., 2014; de Bruin et al., 2022). Considering the significant importance of sialylated and sulfated glycoconjugates in the extracellular matrix of animals, further investigation is needed to see if they also are a common factor in the EPS of microbial aggregates beyond pathogenic microorganisms. Research on their chemical structure and secretion will shed light on their special functionality and evolutionary importance in the extracellular matrix of biofilm in general.

To investigate the presence of sialylated and sulfated glycoconjugates in the EPS of microbial aggregates and to develop specific methodologies to study them, both aerobic and anaerobic granular sludges were collected from full-scale wastewater bioreactors. The alkaline-extracted EPS of these granular sludges was first fractionated by size exclusion chromatography and the collected fractions were analyzed on the presence and diversity of the sialic acids and sulfated glycoconjugates. Genes encoding for known enzymes responsible for the synthesis or transfer of sialic acids and sulfate groups, respectively, were determined by genome database searches on the dominant microorganisms.

2. Materials and Methods

2.1. Experimental set-up

The analysis of the anionic extracellular polymeric substances extracted from both aerobic and anaerobic granular sludge is summarized in Figure 1.

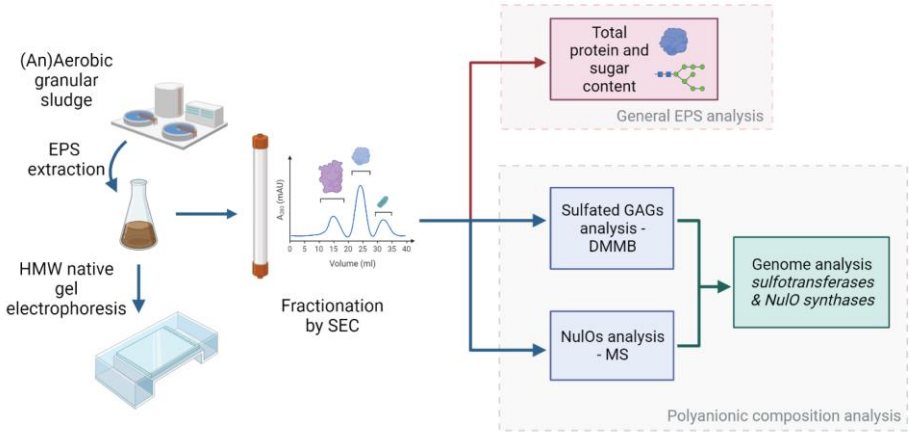


Fig. 1. Schematic representation of the workflow of the analysis of the anionic extracellular polymeric substances. Abbreviations: EPS: extracellular polymeric substances, HMW: High molecular weight, SEC: Size exclusion chromatography, GAGs: glycosaminoglycans, DMMB: 1,9-dimethyl-methylene blue dye, NulOs: nonulosonic acids, MS: mass spectrometry.

2.2 Granular sludge and EPS extraction

The extraction of EPS from aerobic and anaerobic granules were both based on the alkaline-heat extraction method described in previous work (Pinel et al., 2020; Tomás-Martinez et al., 2022; Felz et al., 2016). Aerobic granular sludge was collected from two full-scale wastewater treatment plants (Epe and Zutphen) in the Netherlands, which are operated with the Nereda© Technology. EPS was extracted by alkaline extraction as explained in detail by Bahgat et al. (2023) for the demonstration plants of Epe (sewage) and Zutphen (dairy). The extraction was performed between pH 9 – 11 by the addition of 25% KOH at 80°C. EPS was precipitated afterwards by acidification with 30% HCl to pH 2 – 4. The acid precipitated EPS was dialyzed to retain polymeric components with a 3.5 kDa molecular weight cut off dialysis bag (Snakeskin™, Thermo Fisher Scientific), frozen at -80°C, lyophilized and stored at room temperature until further analysis.

Anaerobic granular sludge was collected from two full-scale anaerobic granular sludge wastewater treatment systems (treating papermill and brewery wastewaters). The EPS was extracted with an alkaline extraction as previously described by Pinel et al. (2020). In short, EPS was extracted by adding dried biomass to 0.1 M NaOH at 80 °C (10 g/L). The sample was stirred vigorously for 30 minutes, after which it was cooled and centrifuged at 3300 $\times g$ for 30 minutes. Supernatant was dialyzed, lyophilized and stored following the same procedure as the EPS from aerobic granular sludge. Detailed information regarding the wastewater treatment process and the type of wastewater is provided in the supplemental material (Table S1).

2.3 Native agarose gel electrophoresis and staining with Coommassie blue and Alcian blue

Native agarose gel electrophoresis was run on a submerged horizontal platform, with the wells positioned in the center of the gel. Lyophilized EPS samples were resolubilized in 50 mM Tris at 2 mg EPS/mL concentration for 1 hour at 30°C. Next, 10 µL of the sample was loaded in the wells on a 0.8% agarose gel in 500 mM Tris/HCl, 160 mM boric acid, 1 M urea, pH 8.5. Electrophoresis was performed with a running buffer (90 mM Tris/HCl, 90 mM boric acid, pH 8.5) at 80 V for 90 min. Proteins carrying a net negative charge migrate toward the anode, whereas proteins carrying a positive charge migrate towards the cathode (Li & Arakawa, 2019). To determine if high molecular weight proteins could pass the gel, a high molecular weight marker was used as a ladder (high molecular weight – SDS Calibration kit, Cytiva, Marlborough, MA). The ladder was negatively charged due to the presence of sodium dodecyl sulfate (SDS). The sample position on the gel was revealed using Coomassie blue staining (SimplyBlue™ Safestain, Invitrogen, Waltham, MA) according to the manufacturer's instruction, and destained in water overnight. To identify the carboxyl groups R-COO^- and the sulfated groups R-OSO_3^-, staining with Alcian blue was performed at pH 2.5 and pH 1.0 respectively, as described by Boleij et al. (2020). The gel pictures were taken on a ChemiDoc MP imager (Bio-Rad, Hercules, CA).

2.4 EPS fractionation by size exclusion chromatography

EPS samples (10 mg) were solubilized in demineralized water to a concentration of 10 mg EPS/mL and the pH was adjusted to 10 using NaOH. All solutions were centrifuged and filtered through a 0.45 µm membrane filter before application to the column, to allow the samples to remain dissolved as much as possible.

Size exclusion chromatography (SEC) was performed using a Hiload 16/600 Superose 6 prepacked column (Cytiva Lifesciences, Marlborough, MA) fitted on a Gilson system containing a UV (280 nm) detector. Calibration of the column, upon which the elution volume was determined, was done using a Cytiva high molecular weight marker set (Cytiva Lifesciences, Marlborough, MA). This consisted of ovalbumin (44 kDa), conalbumin (75 kDa), aldolase (158 kDa), ferritin (440 kDa), thyroglobulin (669 kDa) and blue dextran (2,000 kDa). Blue dextran is usually included to determine the void volume, but Superose 6 has a very high fractionation range (fractionation range M_r ~ 5 kDa-5,000 kDa (globular proteins), exclusion limit M_r ~ 40,000 kDa (globular proteins)), even blue dextran is retained in the column.

Fifteen mL of solubilized EPS samples were run through the column with a flow rate set to 1 mL/min, using a running buffer containing 0.15 M (NaCl) and 0.05 M (glycine) adjusted to pH 10 with NaOH. Five different fractions were chosen based on the retention times of the different proteins in the high molecular weight marker kit and the extrapolation of the calibration line. EPS fractions were subsequently dialyzed to remove excess salts with a 3.5 kDa molecular weight cut-off dialysis bag (Snakeskin™, ThermoFisher Scientific, Landsmeer), frozen at -80°C and lyophilized. The lyophilized samples were stored at room temperature until further analysis.

2.5 Characterization of EPS fractions

2.5.1 Total protein and carbohydrate contents in EPS fractions

Lyophilized EPS fractions were dissolved in 0.01 M NaOH to 0.5 mg/mL. The total protein content was determined by the BCA protein assay following the manufacturer's instruction with bovine serum albumin as a standard (Pierce™ BCA protein assay Kit, Thermo Scientific, USA). Protein absorbance was measured in duplicates at 562 nm using a multimode plate reader (TECAN Infinite M200 PRO, Männedorf, Switzerland). The total carbohydrate content of the EPS solutions, after 2.5 times dilution, was determined by the phenol sulfuric acid method with glucose as a standard (Dubois et al., 1956). The carbohydrate absorbance measurements were performed in cuvettes at 490 nm in duplicates with a VIS-spectrophotometer (HACH DR3900, Ames, IA).

2.5.2 Functional groups of EPS fractions

Functional group analysis was performed by Fourier Transform Infra-red (FT-IR) spectroscopy on a Spectrum 100 spectrometer (PerkinElmer, Shelton, CT). The spectra of the lyophilized samples were recorded at room temperature over a wavenumber range of 600 – 4000 cm^{-1} with 16 accumulations and 4 cm^{-1} resolution.

2.5.3 Sialic acid measurement with mass spectrometry

The NulOs measurement was performed according to the approach described by Kleikamp et al.[7] In short, lyophilized EPS fractions were hydrolyzed by 2 M acetic acid for 2 hours at 80°C and dried with a SpeedVac concentrator. The released NulOs were labelled through the α-keto acid using DMB (1,2-diamino-4,5-methylenedioxybenzene dihydrochloride) for 2.5 hours at 55°C and analyzed by reverse phase chromatography Orbitrap mass spectrometry (QE plus Orbitrap, ThermoFisher Scientific, Bleiswijk, Netherlands). Labelling with other sugars and sugar acids was found to give no DMB derivatives (Hara et al., 1986). To estimate the relative amounts of each type of NulOs, the peak area of 1 μg of KDN was used as a reference signal. The integrated peak areas in the mass spectrometry chromatograms were calculated for each type of sialic acids in each EPS fraction. The peak area was used as a number proportional to the amount of NulOs. The relative amount of each type of sialic acids in each EPS fraction was presented as a ratio to the peak area of 1 μg of KDN for comparison.

2.5.4 Sulfated glycosaminoglycan assay

Detection and quantification of sulfated glycosaminoglycans was performed with the Blyscan sulfated glycosaminoglycan assay (Biocolor, Carrickfergus, UK, assay range is 0-50 ug/mL and the detection limit is 2.5 ug/mL), according to the manufacturer's instructions. Samples (2-5 mg) were digested with one mL of papain protein digestion solution at 65 °C overnight (Sigma Aldrich, Zwijndrecht, Netherlands). The supernatant was recovered after centrifugation at 13000 ×*g* for 10 min. 50 μL of sample was then added to 1 mL of 1,9-dimethyl-methylene blue (DMMB) dye reagent. Sulfated GAGs positive components bind and precipitate with DMMB at a low pH (measured pH in the DMMB solution was 1.7). The precipitate was subsequently isolated and resolubilized. The absorbance of the resolubilized solution at 656 nm (TECAN Infinite M200 PRO, Switzerland) indicated the amount of dye that formed a complex with the sulfated glycosaminoglycans. The standard that was included in the kit was bovine tracheal chondroitin 4-sulfate. Due to the low pH, the influence of intracellular components (*e.g.* DNA) is negligible (Zheng & Levenston, 2015). Lastly, the distribution of N-linked and O-linked sulfate in the samples was measured by performing nitrous acid cleavage as by the manufacturer's instructions prior to sulfated GAGs quantification.

2.6 BLASTp (Protein Basic Local Search Alignment Tool) analysis for nonulosonic acid synthases and sulfotransferases

To identify the 10 most dominant genera of the anaerobic granular sludge community, DNA from sludge samples was extracted using a PowerSoil DNA isolation kit (Qiagen Hilden, Germany) and the V3-V4 regions of the 16S rRNA gene sequenced with primers 341F and 806R (Kleikamp et al., 2022). DNA sequencing was performed at Novogene (Novogene Co., Ltd., China) using Illumina 51 NovaSeq platform. For aerobic granular sludge, the 10 most dominant genera of the community were selected from the study by Kleikamp et al. (2022) BLASTp from the NCBI website was used to identify the homologous enzymes for the biosynthesis of the NulOs and sulfotransferases in close relative organisms (Table

S2) to the most abundant in the anaerobic granular sludge and aerobic granular sludge. The distinct reference protein sequences were taken from bacteria (*Campylobacter jejuni, Bacteroidetes thetheiotaomicron*) or archaea (*Halorubrum* sp PV6) known to produce these types of NulOs. Reference proteins for the NulO synthase of Neu5Ac (NeuB), legionaminic acid (LegI), pseudaminic acid (PseI) and 2-keto-3-deoxynonulosonic acid (KDN-9-phosphate) were used, with corresponding GenBank accession numbers: ERP39285.1, AYD49523, CAL35431 and AAO76821 (Zaretsky et al., 2018; Wang et al., 2008). The distinct reference proteins for sulfotransferases accession numbers: WP_014336261, WP_015887312 (Boleij et al., 2020). Matches with a hit below an *E*-value of 1E-20 were considered significant.

3. Results

EPS was extracted from both anaerobic and aerobic granular sludges with relatively significant amount, i.e. for the two types of anaerobic granular sludge– treating papermill and brewery wastewater – the extraction yield was 43.3 ± 5.5 and 58.4 ± 0.6 %VSS, respectively; for aerobic granular sludge– treating dairy and municipal wastewater the yield was 22.0 ± 1.7 and 29.0 ± 3.1 %VSS, respectively.

3.1 EPS native agarose gel electrophoresis and staining with Coomassie blue and Alcian blue

The extracted EPS were further analyzed by native agarose gel electrophoresis. Following Coomassie Blue staining (Figure 2A), it was observed that, for all EPS samples, a part of the proteins migrated towards the anode (indicative of negatively charged proteins). Another part of the EPS stayed within wells towards the anode, indicating that they may also carry a net negative charge. It is possible that the molecular weight of certain EPS polymers was too high to migrate through the gel (Serwer, 1983).

Alcian blue at pH 2.5 stains both carboxylic and sulfated glycoconjugates, whereas at pH 1.0 it stains only highly negatively charged components, e.g., sulfated glycoconjugates(Boleij et al., 2020; Suvarna et al., 2019). For each EPS, the protein smear (Figure 2A) and the anionic glycoconjugates smear (Figure 2B and 2C) almost overlap with each other. In addition, regarding the part that stays within the well, the pattern stained with Alcian blue at pH 1.0 corresponds to the pattern stained with Coomassie Blue as well. All this information implies that the four EPS samples are all dominated by (glyco)proteins which have carboxylic and sulfated glycoconjugates.

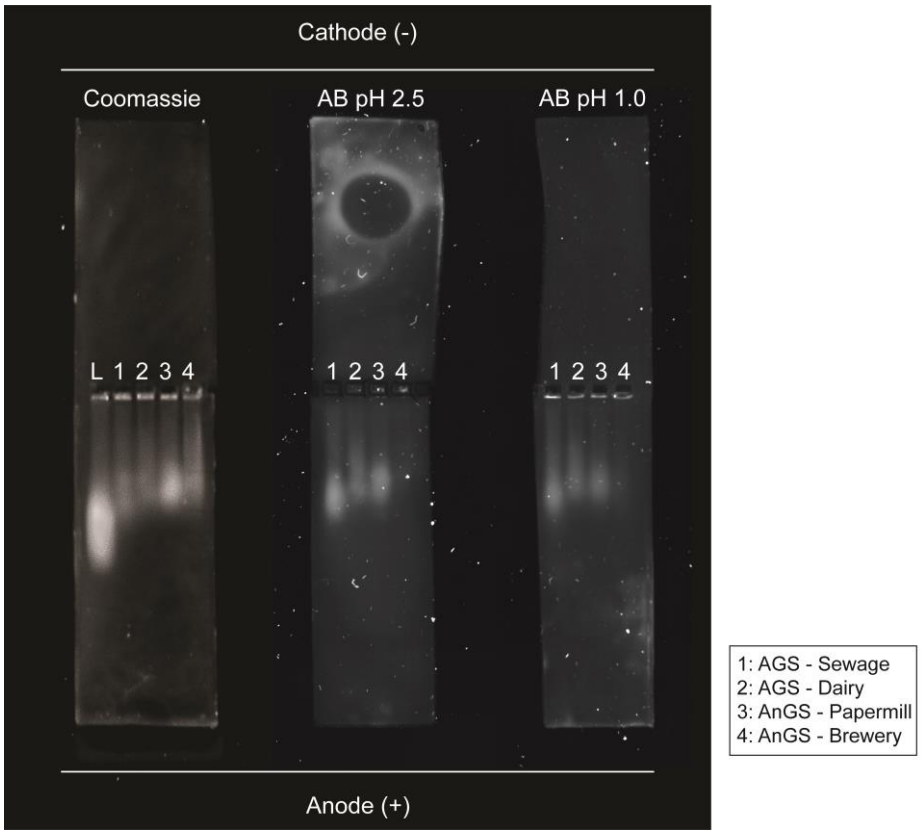


Fig. 2. Native gel electrophoresis on agarose stained with Coomassie G-250 (A), with Alcian blue pH 2.5 (B) and pH 1.0 (C) with the crude EPS from aerobic granular sludge-Sewage (1), aerobic granular sludge-Dairy (2), anaerobic granular sludge-Papermill (3) and anaerobic granular sludge-Brewery and ladder ranging from 53 to 220 kDa (L). The anode (+) is at the bottom of the gels and the cathode (-) is situated at the top.

3.2 EPS fractionation and molecular weight distribution

Native agarose gel electrophoresis indicated that the EPS samples are all dominated by (glyco)proteins and have high molecular weight fractions. To estimate their molecular weight distribution, size exclusion chromatography (SEC) was performed. The detection of proteins' signal at 280 nm was employed to obtain the chromatogram (Edelhoch, 1967). The chromatogram of the EPS samples does not show separate protein peaks, but a continuous curve with absorbance at 280 nm (Figure S1). It is noted that glycosylation of proteins leads to a continuous molecular weight distribution rather than a few specific molecular weights, because level of glycosylation and glycans length can vary for individual proteins (Heijenoort, 2001; Schäffer & Messner, 2017). This observation concurs with the glycosylation of proteins by carboxylic and sulfated glycoconjugates observed with the Alcian blue staining in section 3.2.

The EPS was separated into 5 fractions: 4 fractions in the apparent molecular weight range of 5 kDa to 5,500 kDa, and 1 fraction with an apparent molecular weight >

5,500 kDa. Overall, for each EPS, the mass of the five apparent molecular weight fractions varies (Table 1). Notably, the highest apparent molecular weight fraction (>5,500 kDa) was obtained for every EPS sample and its mass is 16-27% of the mass of the relevant EPS.

Table 1. Fractionation yields for different EPS used after lyophilization (% of fractionated EPS). The non-soluble fraction was not part of the fractionated samples. Samples marked with an asterisk (*) are based on the extrapolation of the calibration line. Actual molecular weights measured would lie between ~2,000 – ~ 40,000 kDa.

		Aerobic granular sludge		Anaerobic granular sludge	
Fraction #	**Molecular weight range (kDa)**	**Sewage**	**Dairy**	**Papermill**	**Brewery**
1	>5,500*	18.8	27.4	17.5	16.9
2	738 – 5,500*	5.6	14.3	7.5	16.2
3	100 - 738	11.2	15.1	13.5	21.8
4	12 - 100	19.6	10.4	29.2	14.6
5	3.5-12	24.7	25.2	30.8	24.1
Non soluble fraction		18.0	7.6	2.4	6.5

3.3 Characterization of the EPS fractions

3.3.1 General EPS characterization – Carbohydrates/ Proteins ratio and functional group analysis

For the fractionated EPS samples, both carbohydrates and proteins were detected in each molecular weight fraction (Figure 3). The sugars to proteins ratio (PS/PN ratio) was significantly higher by 2.7 – 8.6-fold in the highest apparent molecular weight fraction compared to the average of the other fractions. In addition, as the molecular weight decreased, the PS/PN ratio decreased significantly. This indicated that the EPS fractions with apparent molecular weight >5,500 kDa (16.9 - 27.4% by weight of EPS) were probably dominated with glycosylated proteins, while the fractions with apparent molecular weight <5,500 kDa (63.1-81.0% by weight of EPS) were dominated with less or non-glycosylated proteins.

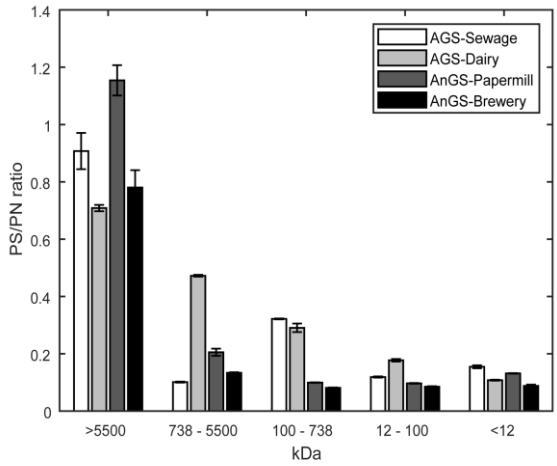


Fig 3. Carbohydrate to protein ratio over the different fractions in aerobic granular sludge-Sewage, aerobic granular sludge-Dairy, anaerobic granular sludge-Papermill and anaerobic granular sludge-Brewery. Carbohydrate content is expressed as glucose equivalents and proteins are expressed as BSA equivalents.

FT-IR spectra were recorded for different EPS fractions to check for the presence of functional groups, *e.g.* amide groups, C-O-C groups for carbohydrates (Figure S2). In addition to the peaks of proteins (1645 cm^{-1} and 1536 cm^{-1}) and carbohydrates (1078 cm^{-1}), two other peaks, which indicate the presence of sialic acids (1730 cm^{-1}) and sulfated esters (1230 cm^{-1}) were observed as well.

3.3.2 Nonulosonic acids

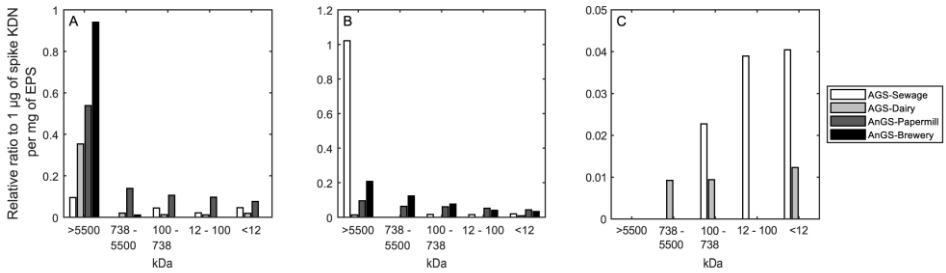


Fig 4. Nonulosonic acids detected in each fraction by MS. The detected NulOs, Leg2Ac/Pse2Ac (A), KDN (B) and NeuAc (C) are expressed as relative ratio of area to spike 1 µg of KDN per mg EPS in each fraction in aerobic granular sludge-Sewage, aerobic granular sludge-Dairy, anaerobic granular sludge-Brewery, anaerobic granular sludge-Papermill

To investigate if sialic acids or other types of NulOs are widespread in different EPS fractions, mass spectrometry analysis was performed. As shown in Figure 4, each EPS fraction was sialylated, with diverse types of NulOs and in different amounts. In total, there were 3 types of NulOs detected: bacterial sialic acids (legionaminic acid (Leg) and/or its stereoisomer pseudaminic acid (Pse)), deaminated neuraminic acid (KDN) and N-acetylneuraminic acid (NeuAc). The most predominant NulO is PseAc2/LegAc2. The highest apparent molecular weight fraction (>5,500 kDa) had the highest amount of PseAc2/LegAc2. Especially for the EPS of the two anaerobic granular sludge, PseAc2/LegAc2 were present to a great extent, 75.5% (brewery) and 99.5% (papermill) of their total amount were located at this fraction. In comparison, the EPS of aerobic granular sludge had slightly lower amount than that of anaerobic granular sludge; about 60.9% (sewage) and 91.3% (dairy) of the total PseAc2/LegAc2 were found in these fractions. The second abundant NulO detected was KDN, except for the EPS from aerobic granular sludge treating dairy wastewater, which had a low amount of KDN compared to the other EPS. The distribution trend was the same as PseAc2/LegAc2: the highest signal of KDN was located at the highest apparent molecular weight fraction (>5,500 kDa). Especially for the EPS from aerobic granular sludge (sewage), 95.2% of the detected KDN was at this fraction. In contrast, the relative amount of NeuAc was on average 16-fold lower than the other two types of NulOs. Only the two aerobic granular sludge EPS had NeuAc which is mainly located at the lower apparent molecular weight fractions (<5,500 kDa).

3.3.3 Sulfated glycosaminoglycans

As FT-IR results indicated the possible presence of sulfate esters, the detection and quantification of sulfate esters such as sulfated glycosaminoglycans was performed. Unlike the profile of NulOs, the presence of sulfated GAGs was widely spread across all samples and sample fractions, with no clear trend (Figure 5). The amount ranged from 7.2 ± 0.1 to 93.7 ± 5.7 μg sulfated GAGs / mg EPS. On average across the molecular weight range, the total sulfated GAGs content in aerobic granular sludge EPS was 64.2 ± 2.2 μg sulfated GAGs / mg EPS and 55.3 ± 1.9 μg sulfated GAGs / mg EPS, for granular sludge from sewage and dairy, respectively. In the anaerobic granular sludge EPS, 42.1 ± 1.6 μg sulfated GAGs / mg EPS and 15.4 ± 0.9μg sulfated GAGs / mg EPS, for granular sludge from papermill and brewery, respectively. In addition to sulfated GAGs, O-linked sulfated GAGs and N-lined sulfated GAGs were determined separately. In the case of aerobic granular sludge, the average weighted percentage of O-linked sulfated GAGs were found to be 46.1 ± 8.6% and 36.6 ± 7.0% in the fractions, for sewage and dairy, respectively. While for anaerobic granular sludge, 29.4 ± 5.8% and 31.9 ± 8.2% O-linked sulfated GAGs were found for papermill and brewery, respectively. Overall, the percentage of O-linked sulfated GAGs was lower than the N-linked sulfated GAGs.

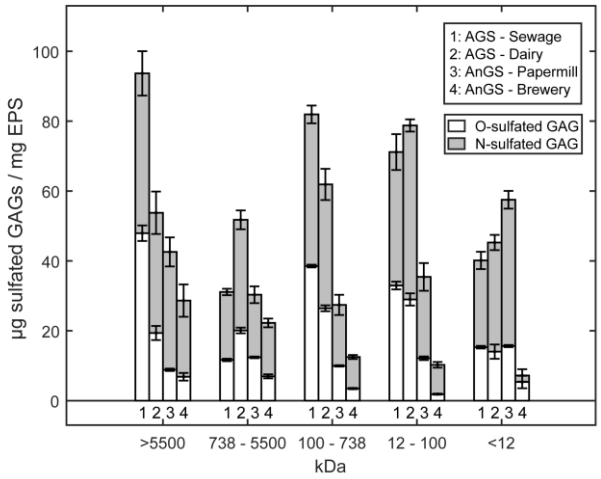


Fig 5. Sulfated glycosaminoglycan concentration (µg sulfated GAGs / mg EPS) of O-linked sulfated GAGs (white) or N-linked sulfated GAGs (grey) detected in each fraction in aerobic granular sludge-Sewage, aerobic granular sludge-Dairy, anaerobic granular sludge-Papermill, anaerobic granular sludge-Brewery.

3.4 Genome analysis of sulfotransferases and nonulosonic acid synthases

To further evaluate the production potential of NulOs and sulfated polymers, BLASTp was performed with key proteins for the formation of these compounds on representative organisms of the top ten most abundant genera in the microbiome of aerobic granular sludge and anaerobic granular sludge (Figure 6). Sulfotransferases, the enzyme which transfers sulfo groups onto polysaccharides, and NulO synthases, the enzyme responsible for the condensation of a 6-carbon sugar with the 3-carbon phosphoenolpyruvate to generate the 9-carbon Leg, Pse, Neu5Ac and KDN, were used as a reference. Most of the genomes from the mined microorganisms contain homologous genes for the NulO biosynthesis of either Neu5Ac, Leg or Pse, implying that these organisms can synthesize NulOs. Hits that matched with the selected NulO synthases were mainly annotated as pseudaminic acid synthase or N-acetylneuraminic acid synthase in the community of aerobic granular sludge and anaerobic granular sludge. A few hits were found for genes annotated with N, N'-diacetyllegionaminic acid synthase (LegI). Differentiation between Neu5Ac and KDN synthases cannot be made due to the way the genes are annotated in the database.

Hits for the sulfotransferases were less abundant. Only few organisms related to representative organisms in the anaerobic granular sludge and aerobic granular sludge microbiome showed positive hits for sulfotransferases. However, when lowering the threshold for the BLASTp search to 5E-2, more organisms showed hits with genes annotated with sulfotransferases. This suggests that the genes encoding for sulfotransferases might have a more distant relation to the reference protein than what is reported for NulOs.

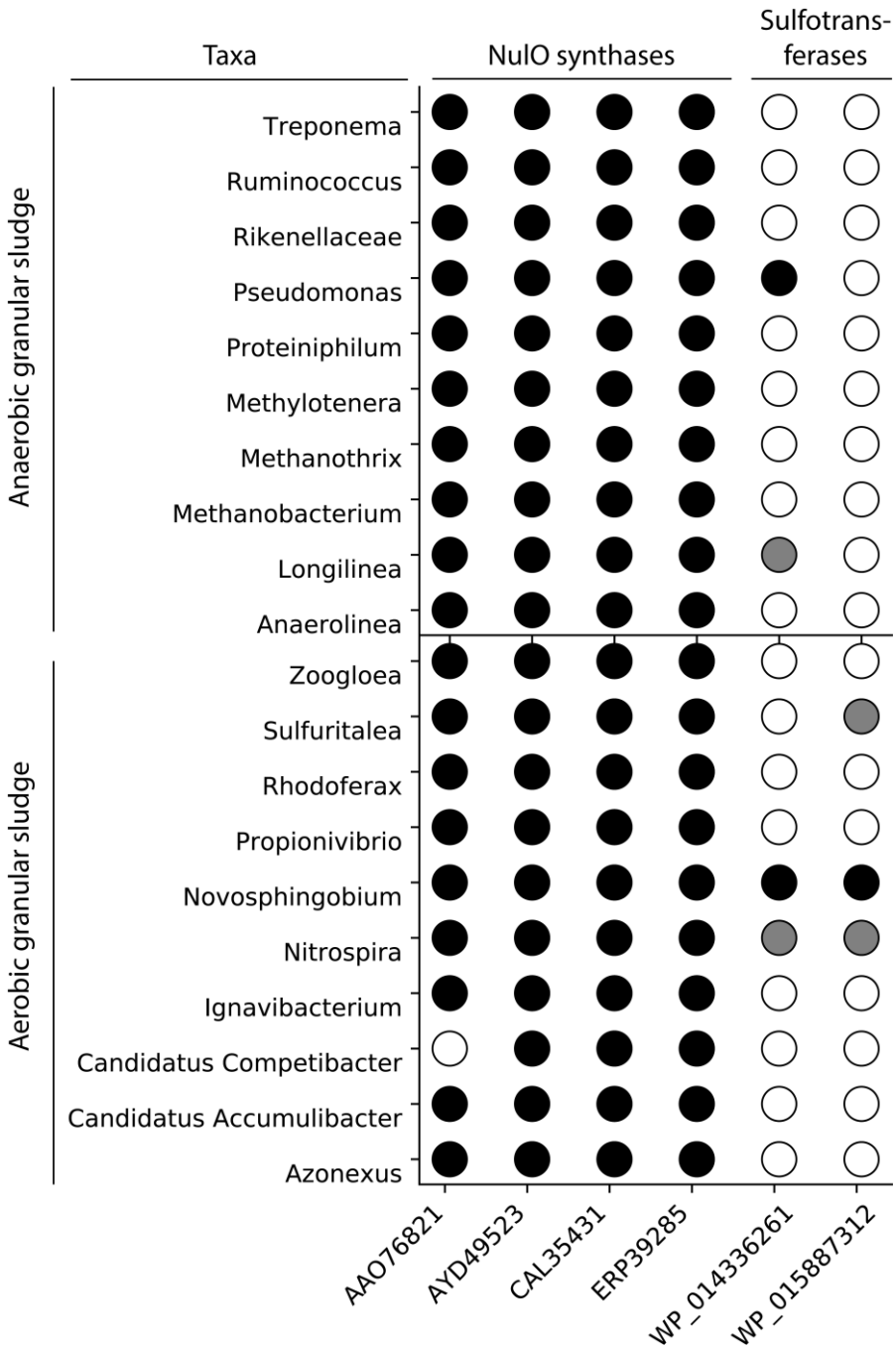


Figure 6. BLASTp analysis of sulfotransferases and the nonulosonic acid synthases over the top 10 most abundant genera in the microbial community of anaerobic granular sludge and aerobic granular sludge. Hits are indicated as a black circle, when the BLASTp analysis showed a match with an *E*-value lower than 1E-20. No hits are indicated as a white circle. Lowering the threshold to 5E-2 revealed more distant hits for sulfotransferases, indicated as a grey circle.

4. Discussion

4.1 Sialylation and sulfation of anionic glycoconjugates are common in the extracellular polymeric substances of both aerobic and anaerobic granular sludge

In the current research, the EPS of anaerobic and aerobic granular sludge collected from two different types of wastewater treatment systems was extracted. The sialylation and sulfation of anionic glycoconjugates in the EPS were investigated. Three types of NulOs were detected, with significant differences in their amount and location: both bacterial sialic acids (PseAc2/LegAc2) and KDN were much more abundant than NeuAc (with the relative amount almost 30 times of NeuAc); the majority of both PseAc2/LegAc2 and KDN were located in the highest molecular weight fraction, while NeuAc was only present in the lower molecular weight fractions of the EPS recovered from aerobic granular sludge. Different from NulOs, the sulfated GAGs were equally distributed over every molecular weight fraction.

The presence of sulfated glycosaminoglycans and NulOs in aerobic granular sludge, anaerobic granular sludge and anammox granular sludge was reported recently (de Graaff et al., 2019; Boleij et al., 2020; Felz et al., 2020; Bourven et al., 2015). Each study focused on either one specific glycoconjugates or one specific sludge. In comparison, in the current research, granular sludges from different waste streams, enriched with different microbial communities (e.g., aerobic and anaerobic microorganisms) and operated under different conditions (e.g., temperature and pH) were investigated. The EPS extraction methods were not identical for the different sludge samples: although both extraction methods are based on harsh alkaline extraction, they varied in scale (i.e., full-scale and lab-scale extraction), the type of the base (i.e. KOH and NaOH), and the subsequent recovery of the solubilized EPS (acidic precipitation and dialysis). Despite of all these differences, it was still observed that NulOs and sulfated groups were present in the EPS samples with similar trends in their abundance and their location at different molecular weight fraction.

Based on the current work and the previous report, the conclusion can be drawn that sialylation and sulfation of anionic glycoconjugates are widely distributed in the extracellular polymeric substances of granular sludge and could be a common phenomenon in environmental biofilms in general. This provides the support that the function of sialylated and sulfated glycoconjugates produced by the microorganisms is not just limited to being a camouflage to avoid the detection of the host immune system, as suggested for pathogenic bacteria, but could be involved in structural components of the granule.

One could speculate on potential functions of sialylated and sulfated glycoconjugates by looking at the role of analogous compounds in animal tissue. Glycosaminoglycans are well-defined polyanionic compounds in the extracellular matrix of animals. These high molecular weight compounds are found widely distributed in the connective tissue, creating a highly porous and hydrophilic hydrogel structure

(Alberts et al., 2002). Furthermore, sialic acids have also been well described in vertebrate cells and are involved in hydration, protein stabilization and cell-cell interactions, due to their negative charge (Varki, 2017). The mass of the highest molecular weight fraction (>5,500 kDa) comprised of 16-27% of the total EPS and was the most glycosylated fraction with a measured carbohydrate content of 35 - 58%. Interestingly, a similar characterization was reported for mucin. Mucins are space-filling large molecular weight glycoproteins (20 – 20,000 kDa) with 50-90% carbohydrate content. The mucin glycoproteins may be sialylated and/or sulfated (Rose et al., 2006; Brown & Hollingworth, 2013). The carbohydrate part are largely involved in the mucin properties, such as hydration, binding of ions and water, and protease inhibitors (Brown & Hollingworth, 2013). It can be speculated that the function of the highest molecular weight EPS fraction might be similar. The exact role of these glycoconjugates inside the extracellular matrix of environmental biofilm is an interesting topic for future investigation.

4.2 Separation and enrichment of the sialylated and sulfated glycoconjugates by size exclusion chromatography with high molecular weight column

The separation and enrichment of glycoconjugates of the extracted EPS aids in further analysis regarding the exact linkage of monomers and chemical structure of these glycoconjugates. This is necessary for a better understanding of the function and production of glycoconjugates in biofilm. Since most sialic acids are on the highest molecular weight fraction (> 5,500 kDa), which is heavily glycosylated and sulfated, enrichment can be achieved at the same time of separation. In this respect, applying a SEC column which can separate large molecular weight polymers is necessary and important. It is noted that, most fractionation studies done with microbial EPS seldom consider protein glycosylation and use columns that can separate molecules up to 670 kDa (Garnier et al., 2006; Simon et al., 2009; Ni et al., 2009). Frequently, the fraction with molecular weight >2,000 kDa is not analyzed since it exceeds the void volume of the column. If the result of the current research is considered, ignoring this fraction leads to the loss of almost 1/3 of the extracted EPS, not to mention the consequence that most of sialylated EPS could never be collected and studied.

High molecular mass biopolymers (molecular weight > 2,000 kDa) are common in nature and are found for instance in the extracellular matrix of vertebrates. Aggrecan, a highly glycosylated and sulfated proteoglycan in articular cartilage, can have molecular mass up to 2,000 kDa (Athanasiou et al., 2017). It has been speculated in literature that bacterial EPS is chemically similar to mucin (Morales-García et al., 2019). Mucin, which is both a sialylated and sulfated glycoprotein complex has a molecular mass of 500 – 3,000 kDa (Alberts et al., 2002; Varki, 2017). Sialylated and/or sulfated glycoconjugates can tremendously increase the molecular mass of proteins.

4.3 Bottlenecks in the study of sialylated and sulfated glycoconjugates in the EPS

Although in the current research, by chemical analysis, both sialylated and sulfated glycoconjugates are found widely spread in all the EPS samples, in the genome analysis of sulfotransferases, very few microorganisms from the most abundant ones in both aerobic granular sludge and anaerobic granular sludge showed positive hits for sulfated glycosaminoglycan sulfotransferases. The reason could be that sulfotransferases were searched in close relatives which do not have this specific gene. Analyzing the metagenome of these samples would improve the estimation of the sulfated glycoconjugates production potential. However, analysis of glycoconjugates production from metagenome data is not trivial. This is due to the fact that metabolic pathways of sulfated GAGs production in bacteria are not well known. The known reference sulfotransferases may not be the enzymes involved in the formation of sulfated glycoconjugates in the EPS. Thus, knowing the exact chemical structure of the sulfated glycoconjugates can aid in finding sulfated polymer production pathways. At present, the methodologies used to study the sulfated glycoconjugates in EPS depend on dye-spectrometric methods, i.e., visualization by Alcian blue staining and heparin red staining, and quantification by DMMB staining (Boleij et al., 2020; Felz et al., 2020; Liu et al., 2021; Bourven et al., 2015). These methods are useful for indicating the presence of sulfated GAGs or other sulfated polymers but have difficulty in distinguishing between different types of sulfated polymers. Sensitive methods like *e.g.*, MS/MS or liquid chromatography (LC)-fluorescence and LC-mass spectrometry (MS) could be used to distinguish between the types of sulfated polymers (Kubaski et al., 2017). As the wide range of molecular size and type of EPS may increase the complexity, the separation of different molecular weight EPS fractions could help with decreasing the complexity and thereby improving the identification of the sulfation patterns.

Separating the EPS by SEC would also help to further determine the exact molecular location of bacterial sialic acids and KDN. Determine the saccharide sequence that the sialic acids are attached to would reveal more information about the structure-function relationship to understand their role in the EPS. Unfortunately, no sequencing tool such as that existing in proteomics or genomics is available to date for glycoconjugates, since glycoconjugates are far more complex and diverse than proteins and nucleic acids (Merritt et al., 2013). In addition, the diversity of NulOs types and modification increase the complexity even more. Altogether, studying NulOs from an environmental sample is challenging. Therefore, to better understand the production and diversity of NulOs in EPS, lectin array or glycoengineering methods can provide novel insights (Ma et al., 2020).

Supporting Information

Information on the wastewater treatment process where samples were collected (supplemental table S1), NCBI tax id and respective number of genomes used in the genome analysis (supplemental table S2), chromatograms of size-exclusion chromatography runs (supplemental figure S1), yields of the subfractions and non-soluble fraction of size exclusion chromatography run (supplemental table S3), absorbance band annotation (supplemental table S4) and FT-IR absorbance spectra for EPS of different molecular weight fractions (supplemental figure S2).

Acknowledgments

This research was financially supported by the SIAM Gravitation Grant 024.002.002, from The Netherlands Organization for Scientific Research. And TKI Chemie 2017 (co-funded by Royal Haskoning DHV), from the Dutch Ministry of Economic Affairs and Climate Policy. The authors thank Martin Pabst and Jitske van Ede for their help with the sialic acid MS analysis. We thank Carina Hof, for help with the size exclusion chromatography analysis.

References

Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. The Extracellular Matrix of Animals. **2002**.

Athanasiou, K. A.; Darling, E. M.; Hu, J. C.; DuRaine, G. D.; Reddi, A. H. Articular Cartilage. **2017**. https://doi.org/10.1201/B14183.

Bahgat, N. T.; Wilfert, P.; Korving, L.; van Loosdrecht, M. Integrated Resource Recovery from Aerobic Granular Sludge Plants. *Water Res.* **2023**, *234*, 119819. https://doi.org/10.1016/J.WATRES.2023.119819.

Boleij, M.; Kleikamp, H.; Pabst, M.; Neu, T. R.; van Loosdrecht, M. C. M.; Lin, Y. Decorating the Anammox House: Sialic Acids and Sulfated Glycosaminoglycans in the Extracellular Polymeric Substances of Anammox Granular Sludge. *Environ. Sci. Technol.* **2020**, acs.est.9b07207. https://doi.org/10.1021/acs.est.9b07207.

Bourven, I.; Bachellerie, G.; Costa, G.; Guibaud, G. Evidence of Glycoproteins and Sulphated Proteoglycan-like Presence in Extracellular Polymeric Substance from Anaerobic Granular Sludge. *http://dx.doi.org/10.1080/09593330.2015.1034186* **2015**, *36* (19), 2428–2435. https://doi.org/10.1080/09593330.2015.1034186.

Brown, R. B.; Hollingsworth, M. A. Mucin Family of Glycoproteins. In *Encyclopedia of Biological Chemistry: Second Edition*; Elsevier Inc., 2013; pp 200–204. https://doi.org/10.1016/B978-0-12-378630-2.00670-8.

Costa, O. Y. A.; Raaijmakers, J. M.; Kuramae, E. E. Microbial Extracellular Polymeric Substances: Ecological Function and Impact on Soil Aggregation. *Front. Microbiol.* **2018**, *9* (JUL), 1636. https://doi.org/10.3389/FMICB.2018.01636/BIBTEX.

de Bruin, S.; Vasquez-Cardenas, D.; Sarbu, S. M.; Meysman, F. J. R.; Sousa, D. Z.; van Loosdrecht, M. C. M.; Lin, Y. Sulfated Glycosaminoglycan-like Polymers Are Present in an Acidophilic Biofilm from a Sulfidic Cave. *Sci. Total Environ.* **2022**, *829*, 154472. https://doi.org/10.1016/J.SCITOTENV.2022.154472.

de Graaff, D. R.; Felz, S.; Neu, T. R.; Pronk, M.; van Loosdrecht, M. C. M.; Lin, Y. Sialic Acids in the Extracellular Polymeric Substances of Seawater-Adapted Aerobic Granular Sludge. *Water Res.* **2019**, 343–351. https://doi.org/10.1016/j.watres.2019.02.040.

Dubois, M.; Gilles, K. A.; Hamilton, J. K.; Rebers, P. A.; Smith, F. Colorimetric Method for Determination of Sugars and Related Substances. *Anal. Chem.* **1956**, *28* (3), 350–356. https://doi.org/10.1021/AC60111A017/ASSET/AC60111A017.FP.PNG_V03.

Edelhoch, H. Spectroscopic Determination of Tryptophan and Tyrosine in Proteins. *Biochemistry* **1967**, *6* (7), 1948–1954. https://doi.org/10.1021/BI00859A010/ASSET/BI00859A010.FP.PNG_V03.

Felz, S.; Al-Zuhairy, S.; Aarstad, O. A.; van Loosdrecht, M. C. M.; Lin, Y. M. Extraction of Structural Extracellular Polymeric Substances from Aerobic Granular Sludge. *J. Vis. Exp.* **2016**, *2016* (115), 54534.

Felz, S.; Neu, T. R.; van Loosdrecht, M. C. M.; Lin, Y. Aerobic Granular Sludge Contains Hyaluronic Acid-like and Sulfated Glycosaminoglycans-like Polymers. *Water Res.* **2020**, *169*. https://doi.org/10.1016/j.watres.2019.1

15291.

Flemming, H. C.; Wingender, J. The Biofilm Matrix. *Nature Reviews Microbiology*. Nature Publishing Group September 2, 2010, pp 623–633. https://doi.org/10.1038/nrmicro2415.

Garnier, C.; Gorner, T.; Guinot-Thomas, P.; Chappe, P.; de Donato, P. Exopolymeric Production by Bacterial Strains Isolated from Activated Sludge of Paper Industry. *Water Res.* **2006**, *40* (16), 3115–3122. https://doi.org/10.1016/J.WATRES.2006.06.005.

Hara, S.; Yamaguchi, M.; Takemori, Y.; Nakamura, M.; Ohkura, Y. Highly Sensitive Determination of N-Acetyl- and N-Glycolylneuraminic Acids in Human Serum and Urine and Rat Serum by Reversed-Phase Liquid Chromatography with Fluorescence Detection. *J. Chromatogr. B Biomed. Sci. Appl.* **1986**, *377* (C), 111–119. https://doi.org/10.1016/S0378-4347(00)80766-5.

Heijenoort, J. van. Formation of the Glycan Chains in the Synthesis of Bacterial Peptidoglycan. *Glycobiology* **2001**, *11* (3), 25R-36R. https://doi.org/10.1093/GLYCOB/11.3.25R. https://doi.org/10.3791/54534.

Kleikamp, H. B. C.; Grouzdev, D.; Schaasberg, P.; van Valderen, R.; van der Zwaan, R.; van de Wijgaart, R.; Lin, Y.; Abbas, B.; Pronk, M.; van Loosdrecht, M. C. M.; Pabst, M. Comparative Metaproteomics Demonstrates Different Views on the Complex Granular Sludge Microbiome. *bioRxiv* **2022**, 2022.03.07.483319. https://doi.org/10.1101/2022.03.07.483319.

Kleikamp, H. B. C.; Lin, Y. M.; McMillan, D. G. G.; Geelhoed, J. S.; Naus-Wiezer, S. N. H.; Van Baarlen, P.; Saha, C.; Louwen, R.; Sorokin, D. Y.; Van Loosdrecht, M. C. M.; Pabst, M. Tackling the Chemical Diversity of Microbial Nonulosonic Acids-a Universal Large-Scale Survey Approach. *Chemical Science*. Royal Society of Chemistry March 21, 2020, pp 3074–3080. https://doi.org/10.1039/c9sc06406k.

Kubaski, F.; Osago, H.; Mason, R. W.; Yamaguchi, S.; Kobayashi, H.; Tsuchiya, M.; Orii, T.; Tomatsu, S. Glycosaminoglycans Detection Methods: Applications of Mass Spectrometry. *Mol. Genet. Metab.* **2017**, *120* (1–2), 67–77. https://doi.org/10.1016/J.YMGME.2016.09.005.

Li, C.; Arakawa, T. Agarose Native Gel Electrophoresis of Proteins. *Int. J. Biol. Macromol.* **2019**, *140*, 668–671. https://doi.org/10.1016/J.IJBIOMAC.2019.08.066.

Liu, J.; Zhang, Z.; Xue, W.; Siriweera, W. B.; Chen, G.; Wu, D. Anaerobic Digestion of Saline Waste Activated Sludge and Recovering Raw Sulfated Polysaccharides. *Bioresour. Technol.* **2021**, *335*, 125255. https://doi.org/10.1016/J.BIORTECH.2021.125255.

Ma, B.; Guan, X.; Li, Y.; Shang, S.; Li, J.; Tan, Z. Protein Glycoengineering: An Approach for Improving Protein Properties. *Front. Chem.* **2020**, *8*, 622. https://doi.org/10.3389/FCHEM.2020.00622/BIBTEX.

Masuko, T.; Minami, A.; Iwasaki, N.; Majima, T.; Nishimura, S. I.; Lee, Y. C. Carbohydrate Analysis by a Phenol-Sulfuric Acid Method in Microplate Format. *Anal. Biochem.* **2005**, *339* (1), 69–72. https://doi.org/10.1016/J.AB.2004.12.001.

Merritt, J. H.; Ollis, A. A.; Fisher, A. C.; Delisa, M. P. Glycans-by-Design: Engineering Bacteria for the Biosynthesis of Complex Glycans and Glycoconjugates. *Biotechnol. Bioeng.* **2013**, *110* (6), 1550–1564. https://doi.org/10.1002/BIT.24885.

Morales-García, A. L.; Bailey, R. G.; Jana, S.; Burgess, J. G. The Role of Polymers in Cross-Kingdom Bioadhesion. *Philos. Trans. R. Soc. B* **2019**, *374* (1784). https://doi.org/10.1098/RSTB.2019.0192.

Morrison, M. J.; Imperiali, B. The Renaissance of Bacillosamine and Its Derivatives:Pathway Characterization and Implications in Pathogenicity. *Biochemistry* **2014**, *53* (4), 624. https://doi.org/10.1021/BI401546R.

Ni, B. J.; Fang, F.; Xie, W. M.; Sun, M.; Sheng, G. P.; Li, W. H.; Yu, H. Q. Characterization of Extracellular Polymeric Substances Produced by Mixed Microorganisms in Activated Sludge with Gel-Permeating Chromatography, Excitation–Emission Matrix Fluorescence Spectroscopy Measurement and Kinetic Modeling. *Water Res.* **2009**, *43* (5), 1350–1358. https://doi.org/10.1016/J.WATRES.2008.12.004.

Parc, A. Le; Lee, H.; Chen, K.; Barile, D.; Parc, A. Le; Lee, H.; Chen, K.; Barile, D. Rapid Quantification of Functional Carbohydrates in Food Products. *Food Nutr. Sci.* **2014**, *5* (1), 71–78. https://doi.org/10.4236/FNS.2014.51010.

Pinel, I. S. M.; Kleikamp, H. B. C.; Pabst, M.; Vrouwenvelder, J. S.; van Loosdrecht, M. C. M.; Lin, Y. Sialic Acids: An Important Family of Carbohydrates Overlooked in Environmental Biofilms. *Appl. Sci. 2020, Vol. 10, Page 7694* **2020**, *10* (21), 7694. https://doi.org/10.3390/APP10217694.

Pronk, M.; de Kreuk, M. K.; de Bruin, B.; Kamminga, P.; Kleerebezem, R.; van Loosdrecht, M. C. M. Full Scale Performance of the Aerobic Granular Sludge Process for Sewage Treatment. *Water Res.* **2015**, *84*, 207–217. https://doi.org/10.1016/j.watres.2015.07.011.

Rose, M. C.; Voynow, J. A. Respiratory Tract Mucin Genes and Mucin Glycoproteins in Health and Disease. *Physiol. Rev.* **2006**, *86* (1), 245–278. https://doi.org/10.1152/PHYSREV.00010.2005/ASSET/IMAGES/LARGE/Z9J0010623871006.JPEG.

Schäffer, C.; Messner, P. Emerging Facets of Prokaryotic Glycosylation. *FEMS Microbiol. Rev.* **2017**, *41* (1), 49. https://doi.org/10.1093/FEMSRE/FUW036.

Serwer, P. Agarose Gels: Properties and Use for Electrophoresis. *Electrophoresis* **1983**, *4* (6), 375–382. https://doi.org/10.1002/ELPS.1150040602.

Simon, S.; Païro, B.; Villain, M.; D'Abzac, P.; Hullebusch, E. Van; Lens, P.; Guibaud, G. Evaluation of Size Exclusion Chromatography (SEC) for the Characterization of Extracellular Polymeric Substances (EPS) in Anaerobic Granular Sludges. *Bioresour. Technol.* **2009**, *100* (24), 6258–6268. https://doi.org/10.1016/j.biortech.2009.07.013.

Suvarna, S. K.; Layton, C.; Bancroft, J. D. Bancroft's THEORY and TECHNIQUES of HISTOLOGICAL PRACTICE. *Anal. Standar Pelayanan Minimal Pada Instal. Rawat Jalan di RSUD Kota Semarang* **2019**, *3*, 557.

Tomás-Martínez, S.; Chen, L. M.; Pabst,

M.; Weissbrodt, D. G.; van Loosdrecht, M. C. M.; Lin, Y. Enrichment and Application of Extracellular Nonulosonic Acids Containing Polymers of Accumulibacter. *Appl. Microbiol. Biotechnol.* **2022**, 1–11.

Tomás-Martínez, S.; Kleikamp, H. B. C.; Neu, T. R.; Pabst, M.; Weissbrodt, D. G.; van Loosdrecht, M. C. M.; Lin, Y. Production of Nonulosonic Acids in the Extracellular Polymeric Substances of "Candidatus Accumulibacter Phosphatis." *Appl. Microbiol. Biotechnol.* **2021**, *105* (8), 3327–3338. https://doi.org/10.1007/s00253-021-11249-3.

van Lier, J. B.; Mahmoud, N.; Zeeman, G. Anaerobic Biological Wastewater Treatment. In *Biological wastewater treatment*; Henze, M., van Loosdrecht, M., Ekama, G. A., Brdjanovic, D., Eds.; IWA publishing, 2008; pp 415–457.
Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3–49. https://doi.org/10.1093/GLYCOB/CWW086.

Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H.; Schnaar, R. L.; Seeberger, P. H. Essentials of Glycobiology. *Cold Spring Harb.* **2017**, 823.

Wang, L.; Lu, Z.; Allen, K. N.; Mariano, P. S.; Dunaway-Mariano, D. Human Symbiont Bacteroides Thetaiotaomicron Synthesizes 2-Keto-3-Deoxy-D-Glycero-D- Galacto-Nononic Acid (KDN). *Chem. Biol.* **2008**, *15* (9), 893–897. https://doi.org/10.1016/j.chembiol.2008.08.005.

Zaretsky, M.; Roine, E.; Eichler, J. Sialic Acid-like Sugars in Archaea: Legionaminic Acid Biosynthesis in the Halophile Halorubrum Sp. PV6. *Front. Microbiol.* **2018**, *9* (SEP), 2133. https://doi.org/10.3389/FMICB.2018.02133/BIBTEX.

Zheng, C. H.; Levenston, M. E. Fact versus Artifact: Avoiding Erroneous Estimates of Sulfated Glycosaminoglycan Content Using the Dimethylmethylene Blue Colorimetric Assay for Tissue-Engineered Constructs. *Eur. Cells Mater.* **2015**, *29*, 224–236. https://doi.org/10.22203/eCM.v029a17.

Supplemental Information

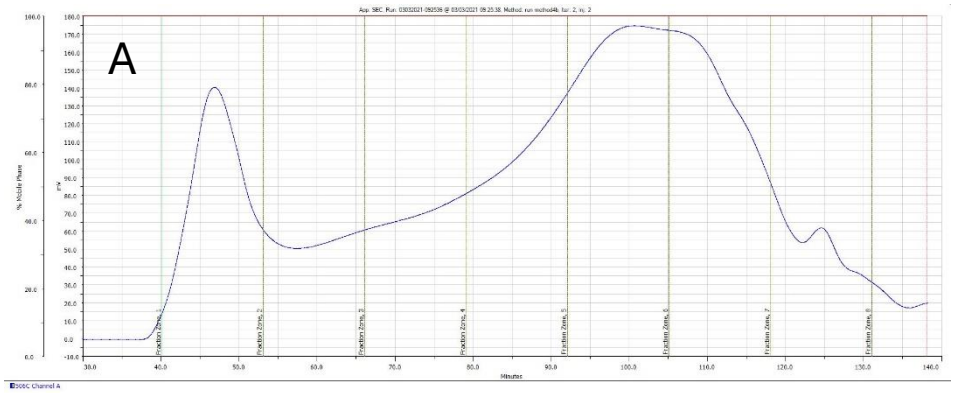
Supplemental Table S1. Information of the aerobic and anaerobic wastewater treatment process where the sludge was collected.

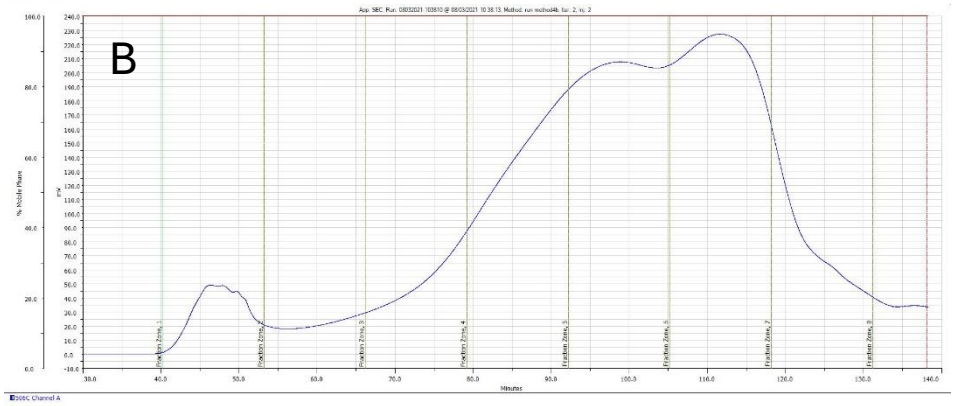
	AGS-Sewage	**AGS-Dairy**	**AnAGS-Brw**	**AnAGS-Pap**
wastewater treatment process	Nereda© Technology*	Nereda© Technology*	UASB	IC
Wastewater type	75% municipal 25% slaughterhouse	Industrial wastewater from dairy factory	Industrial wastewater from brewery	Industrial wastewater from paper industry (recycle and kraft)

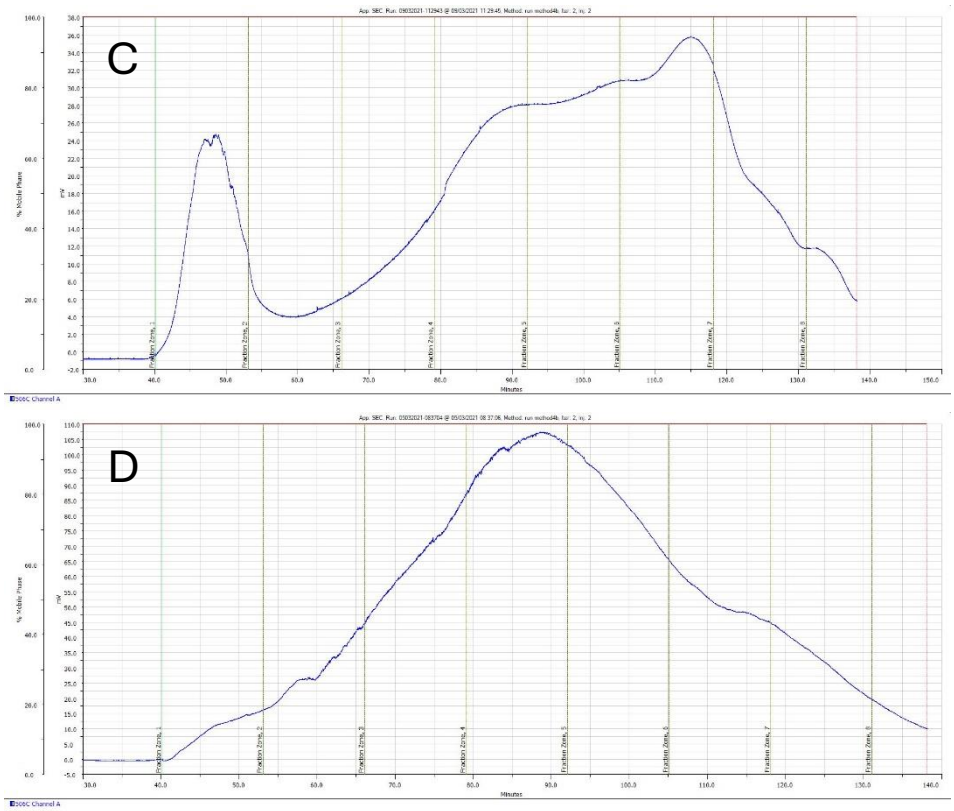
*Nereda® is a registered trademark for an aerobic granular sludge technology owned by Royal HaskoningDHV (Pronk et al (2015). Full scale performance of the aerobic granular sludge process for sewage treatment. Water Research, 84, 207–217. https://doi.org/10.1016/j.watres.2015.07.011)

Supplementary Table S2. NCBI tax id and their respective number of genomes of the microorganisms used in the genome analysis.

GS type	Name organism	NCBI:txid	Number of genomes
AnGS	Treponema	157	1240
	Ruminococcus	1263	2,149
	Rikenellaceae	171550	1,620
	Pseudomonas	286	26,150
	Proteiniphilum	294702	62
	Methylotenera	359407	80
	Methanothrix	2222	113
	Methanobacterium	2160	120
	Longilinea	475961	5
	Anaerolinea	233189	49
AGS	Zoogloea	349	29
	Sulfuritalea	1054211	32
	Rhodoferax	28065	146
	Propionivibrio	83766	28
	Novosphingobium	165696	246
	Nitrospira	1234	299
	Ignavibacterium	795750	33
	Candidatus Competibacter	221279	6
	Candidatus Accumulibacter	327159	82
	Azonexus	146936	20


A
% Nickle Phase
Minutes


B
% Nickle Phase
Minutes



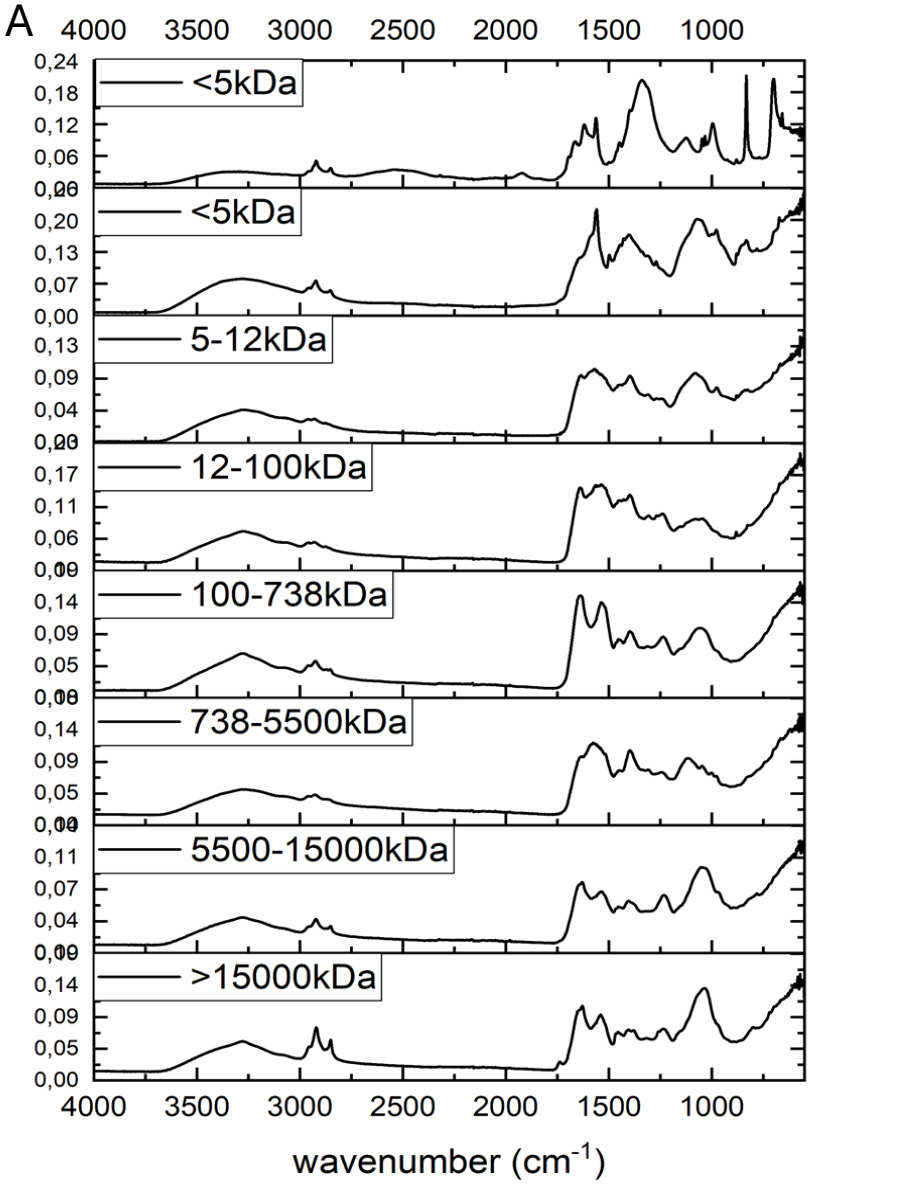
Supplemental Figure S1. Chromatogram of SEC runs at 280 nm of the extracted EPS of AGS-Sewage (A), AGS-Dairy (B), AnGS-Papermill (C) and AnGS-Brewery (D)

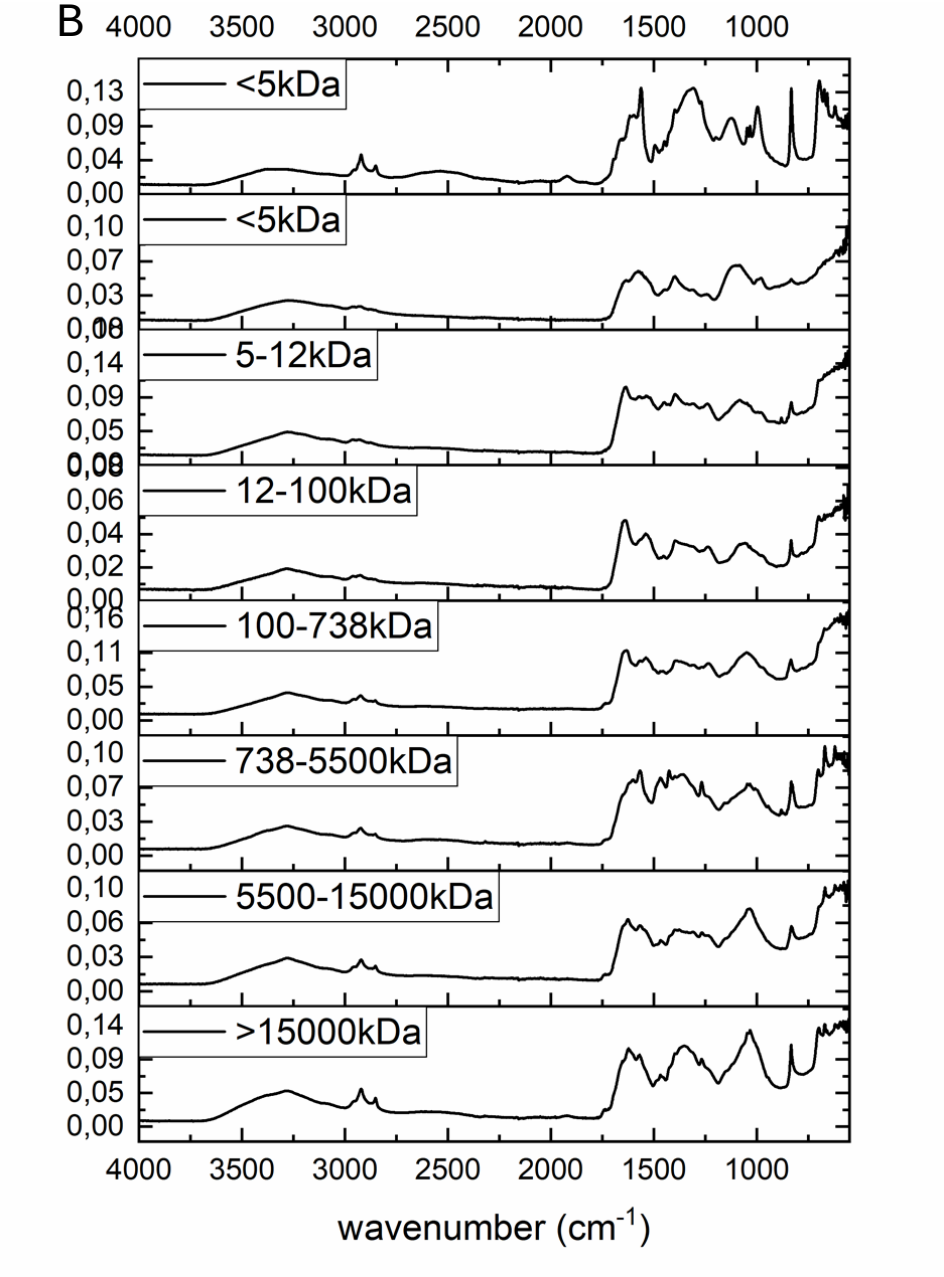
Supplemental Table S3. Subfractions yields for different EPS used after lyophilization (% of fractionated EPS). The non-soluble fraction was not part of the fractionated samples.

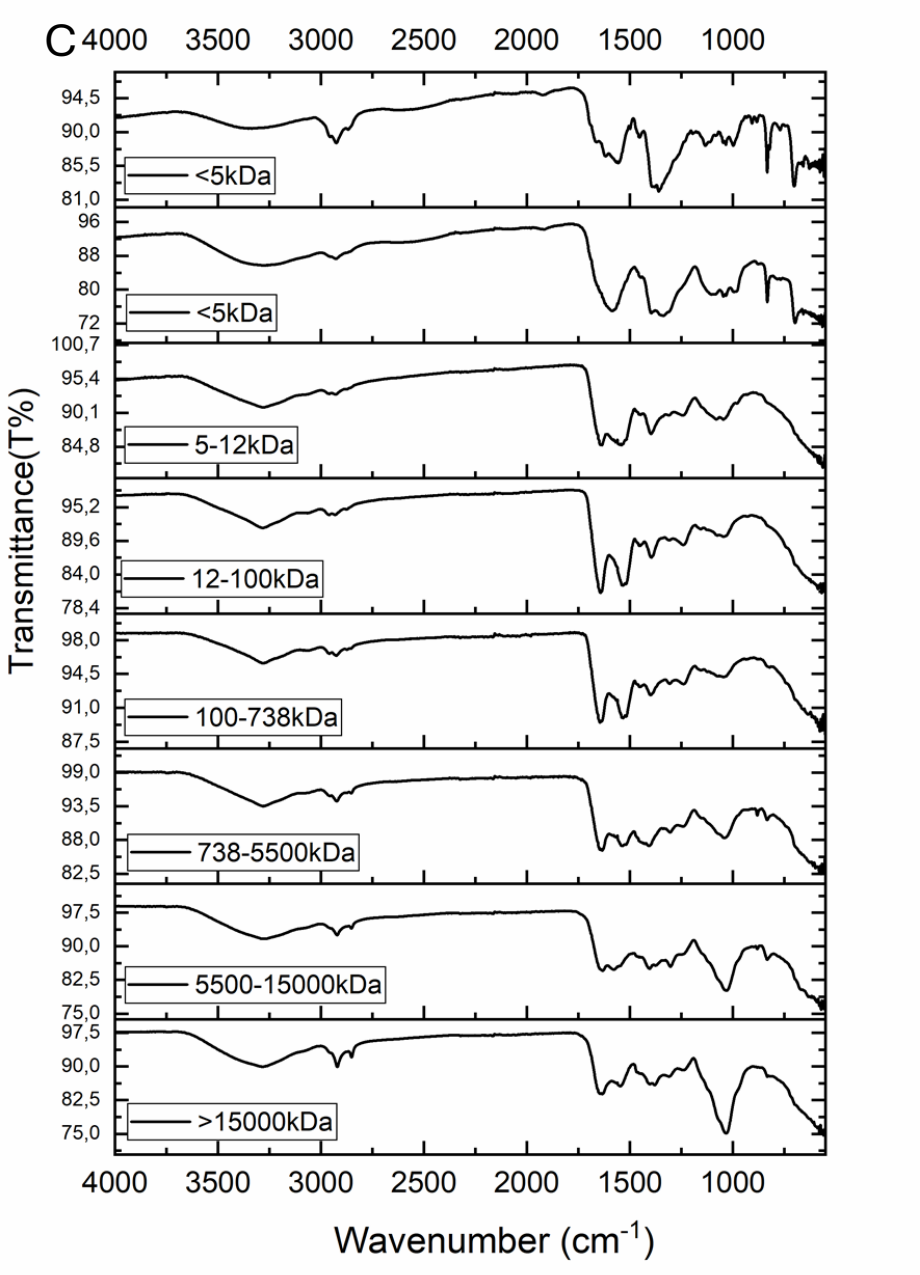
fraction #	MW range (kDa)	AGS-Sewage (% of fractionated EPS)	AGS-Dairy (% of fractionated EPS)	AnGS-Papermill (% of fractionated EPS)	AnGS-Brewery(% of fractionated EPS)
1	>15,000	11.8	14.9	10.1	6.7
2	5,500–15,000	7.0	12.5	7.4	10.2
3	738 – 5,500	5.6	14.3	7.5	16.2
4	100 - 738	11.2	15.1	13.5	21.8
5	12 - 100	19.6	10.4	29.2	14.6
6	5 - 12	18.6	15.5	22.4	14.3
7	3 - 5	8.1	9.7	8.4	9.8
Non soluble fraction		18.0	7.6	2.4	6.5

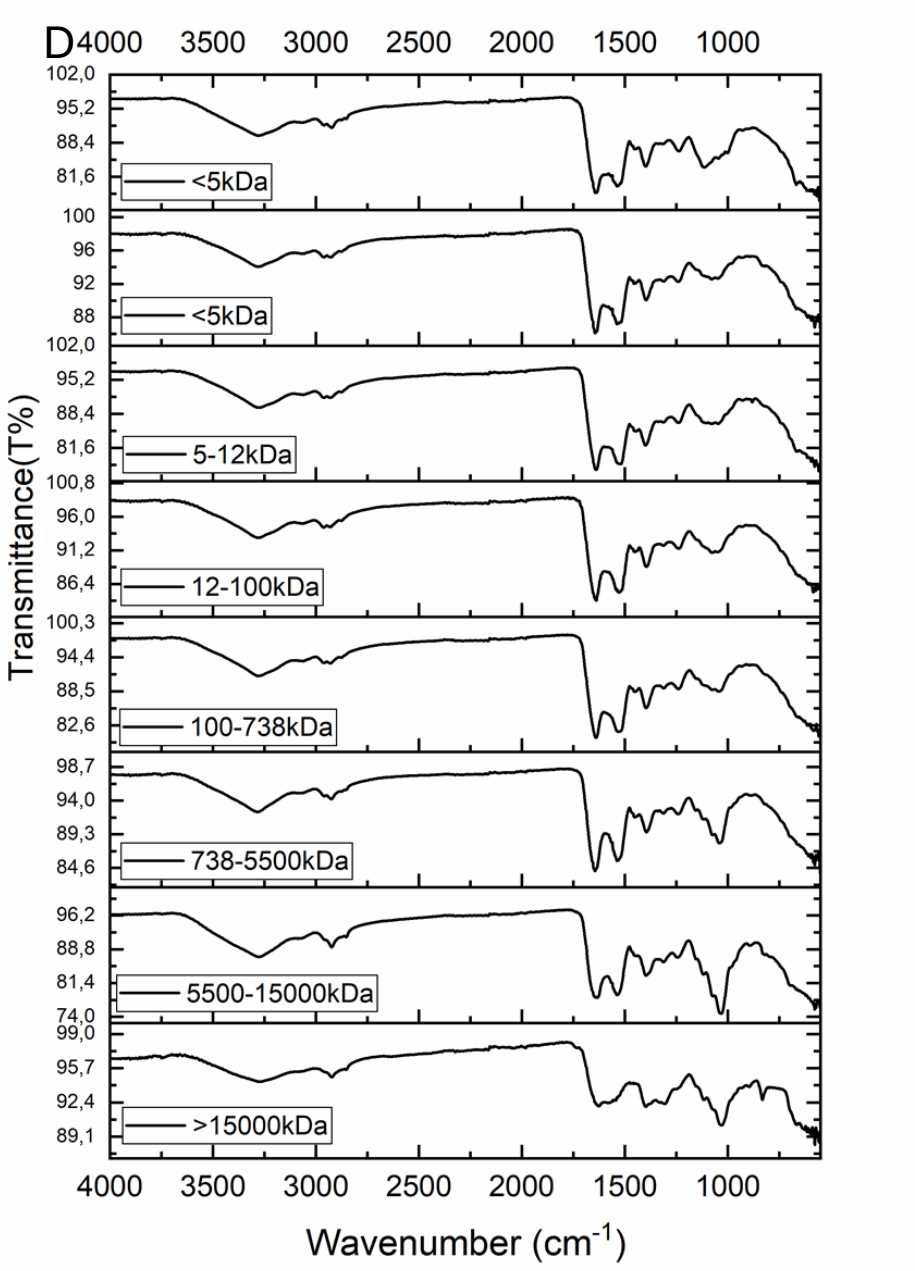
Supplemental Table S4. Functional group assignment of absorbance in specific wavenumbers.

Wavenumber	Functional group assignment	reference
1730 cm-1	ν(C=O)OH stretch sialic acid	de Graaff et al., 2019
1645 cm-1	Protein Amide I	Talari et al., 2017; Kong et al., 2007
1536 cm-1	Protein Amide II	Talari et al., 2017; Kong et al., 2007
1230 cm-1	ν_{as} S=O stretch of sulfate ester	Cabassi et al., 1978
1078 cm-1	-CO (β-glycosidic) carbohydrates	Ma et al., 2018
1025 cm-1	-CO (α-glycosidic) carbohydrates	Ma et al., 2018


A
4000
3500
3000
2500
2000
1500
1000
<5kDa
<5kDa
5-12kDa
12-100kDa
100-738kDa
738-5500kDa
5500-15000kDa
>15000kDa
wavenumber (cm-1)


B
4000 3500 3000 2500 2000 1500 1000
<5kDa
0,13
0,09
0,04
0,00
<5kDa
0,10
0,07
0,03
5-12kDa
0,14
0,09
0,05
12-100kDa
0,06
0,04
0,02
0,00
100-738kDa
0,16
0,11
0,05
0,00
738-5500kDa
0,10
0,07
0,03
0,00
5500-15000kDa
0,10
0,06
0,03
0,00
>15000kDa
0,14
0,09
0,05
0,00
4000 3500 3000 2500 2000 1500 1000
wavenumber (cm-1)


C
4000
3500
3000
2500
2000
1500
1000
94,5
90,0
85,5
81,0
<5kDa
96
88
80
72
<5kDa
100,7
95,4
90,1
84,8
5-12kDa
95,2
89,6
84,0
78,4
12-100kDa
98,0
94,5
91,0
87,5
100-738kDa
99,0
93,5
88,0
82,5
738-5500kDa
97,5
90,0
82,5
75,0
5500-15000kDa
97,5
90,0
82,5
75,0
>15000kDa
Transmittance(T%)
Wavenumber (cm-1)



Supplementary Figure S2. FTIR Spectra of the EPS fractions and subfractions of AGS-Sewage (A), AGS-Dairy (B), AnGS-Papermill (C) and AnGS-Brewery (D). The reference wavelengths were identified to its according functional group (Table S4).

Additional references

Talari, A. C. S., Martinez, M. A. G., Movasaghi, Z., Rehman, S., & Rehman, I. U. (2017). Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews*, *52*(5), 456-506.

Kong, J., & Yu, S. (2007). Fourier transform infrared spectroscopic analysis of protein secondary structures. *Acta biochimica et biophysica Sinica*, *39*(8), 549-559.

Cabassi, F., Casu, B., & Perlin, A. S. (1978). Infrared absorption and Raman scattering of sulfate groups of heparin and related glycosaminoglycans in aqueous solution. *Carbohydrate Research*, *63*, 1-11.

Ma, Y., He, H., Wu, J., Wang, C., Chao, K., & Huang, Q. (2018). Assessment of polysaccharides from mycelia of genus Ganoderma by mid-infrared and near-infrared spectroscopy. *Scientific Reports*, *8*(1), 1-1

Chapter 4:
Coupling extracellular glycan composition with metagenomic data in papermill and brewery anaerobic granular sludges

Abstract
Glycans are crucial for the structure and function of anaerobic granular sludge in wastewater treatment. Yet, there is limited knowledge regarding the microorganisms and biosynthesis pathways responsible for glycan production. In this study, we analysed samples from anaerobic granular sludges treating papermill and brewery wastewater, examining glycans composition and using metagenome-assembled genomes (MAGs) to explore potential biochemical pathways associated with their production. Uronic acids were the predominant constituents of the glycans in extracellular polymeric substances (EPS) produced by the anaerobic granular sludges, comprising up to 60% of the total polysaccharide content. MAGs affiliated with *Anaerolineacae, Methanobacteriaceae* and *Methanosaetaceae* represented the majority of the microbial community (30-50% of total reads per MAG). Based on the analysis of MAGs, it appears that *Anaerolinea* sp. and members of the *Methanobacteria* class are involved in the production of exopolysaccharides within the analysed granular sludges. These findings shed light on the functional roles of microorganisms in glycan production in industrial anaerobic wastewater treatment systems.

Keywords
Anaerobic granular sludge, glycans, metagenomics, extracellular polymeric substances, EPS biosynthesis pathways

Highlights

- EPS of papermill and brewery anaerobic granular sludges have similar glycan composition.
- Sugar monomers of the studied granular EPS had 40 % (w/w) galacturonic acid content.
- MAGs of *Anaerolineacae* and *Methanobacteriaceae* were most abundant in the granular sludges.
- Abundant MAGs had the most genes for the biosynthesis of UDP-N-acetyl-D-galactosamine.

Published as:

Doloman, A., de Bruin, S., van Loosdrecht, M. C., Sousa, D. Z., & Lin, Y. (2024). Coupling extracellular glycan composition with metagenomic data in papermill and brewery anaerobic granular sludges. Water Research, 121240.

1. Introduction

Anaerobic wastewater treatment is a widely applied technology to convert industrial wastewater into methane-rich biogas, a renewable energy carrier (van Lier et al., 2015a). This technology relies on the activity of microbial communities, often self-assembled into aggregates (granules) of 1-5 mm diameter. Sludge granulation is crucial for compact high-rate reactor technology, as the high density of granules permits the uncoupling of solid and hydraulic retention times (Hulshoff Pol et al., 2004; van Lier et al., 2015b).

The phenomenon of sludge granulation is linked to the production of extracellular polymeric substances (EPS) by microorganisms (Schmidt and Ahring, 2000). These EPS form a matrix that embeds the microorganisms in a gel-like structure, providing protection against external stressors. Previous studies on surface-associated biofilm-forming microorganisms revealed the diverse composition of EPS, including proteins, polysaccharides and lipids, and their combinations, like glycoproteins and lipopolysaccharides (Neu and Kuhlicke, 2017; Seviour et al., 2019). The ability of the extracellular matrix to form hydrogels is attributed mainly to the presence of functional groups in EPS that carry negative charges (Decho and Gutierrez, 2017), e.g. sialic acids, sulfated polymers and charged sugar monomers (Boleij et al., 2020; de Bruin et al., 2022; Felz et al., 2020). Yet, due to the complexity of glycans present in EPS, the composition of EPS individual glycans in the glycome (polysaccharides and glycoconjugates) remains relatively unexplored (Seviour et al., 2019). This is particularly the case in the context of mixed communities like anaerobic granular sludge. There, specific sugar monomers within glycans, such as rhamnose, mannose, glucose and galactosamine, are believed to be significant (Veiga et al., 1997). Furthermore, research also indicates that the glycoconjugate composition in anaerobic granular sludge is highly dynamic and diverse, allowing microorganisms to adapt swiftly to their environment. For example, anaerobic granules adapted to saline wastewater (20 g/L Na^+) showed an elevated content of N-acetyl-galactosamine and galactose glycoconjugates compared to granules grown in low-salinity (5 g/L Na^+) wastewater (Gagliano et al., 2018).

Related to the lack of knowledge on the glycome composition of anaerobic biofilms, the EPS synthesis pathways also remain poorly characterized in anaerobes. On the contrary, in aerobes several pathways have been described for EPS synthesis, including the wzy/wzx-dependent, the synthase-dependent, and the ABC transporter-dependent pathways (Limoli et al., 2015; Sun and Zhang, 2021; Whitfield et al., 2020). Genes required for EPS production encode functions related to regulation, synthesis of sugar monomers, chain-length determination, repeat-unit assembly, polymerization and export (Schmid et al., 2015). However, these genes are not always grouped in operons, and some may overlap with central metabolism, making glycoconjugates and exopolysaccharide-associated biosynthetic gene clusters challenging to detect (Jennings et al., 2015). A recent large-scale metagenomic study of aerobic sludge found that 11% of the metagenome assembled genomes (MAGs) contained known operons with all necessary genes to produce extracellular polysaccharides (Dueholm et al., 2023). This indicates that either EPS biosynthesis

potential is restricted to a small subset of the microbial community, or that new EPS biosynthetic clusters remain to be discovered.

This study aimed to provide a way to qualitatively link the EPS glycome of anaerobic granular sludge with the microbial glycan-biosynthetic potential. By linking the chemical and metagenomic data it would be possible to guide the search of novel EPS biosynthetic clusters. Samples of anaerobic granular sludge were collected from industrial bioreactors with two distinct configurations (IC and UASB) and treating wastewater of varying composition (papermill or brewery wastewater). Through a combination of chemical analyses of the sludge EPS glycome and metagenomic sequencing, it was possible to study the glycan and microbial composition of the anaerobic granular sludges. Additionally, by analyzing MAGs encoding biosynthesis genes for the glycan-associated sugar monomers, it was possible to pinpoint potential microorganisms contributing to the production of the anaerobic sludge EPS glycome.

2. Materials and methods

2.1. Source of anaerobic granular sludge

Sludge samples were obtained from three full-scale anaerobic digesters, each with specific configuration and wastewater source (Table 1). Two internal circulation (IC) reactors received wastewater from papermill industry, while the upflow anaerobic sludge blanket (UASB) reactor was supplied with brewery wastewater. Industrial paper recycle/kraft wastewater can be highly contaminated with lignin, lignin-derivatives, and toxic compounds that are difficult to digest (Kumar et al., 2022). Brewery wastewater is characterized by a high organic content derived from sugars, soluble starch, and ethanol (Simate et al., 2011).

Table 1. Anaerobic granular sludge characteristics.

Sample	Papermill A	Papermill B	Brewery
Reactor type	IC	IC	UASB
Wastewater source	Paper recycle & kraft	Paper recycle & kraft	Brewery
Temperature (°C)	36	35	29
Influent COD (mg/L)	1200	1200	3900
Effluent VFA (mEq/L)	124.0	112.2	47.2
Pre-acidification (% of COD)	19 %	19 %	n.a.
Granule morphology	Round, slightly fluffy	Round, smooth	Round, grainy
Color	beige	beige	black

COD – chemical oxygen demand; VFA – volatile fatty-acids; pre-acidification – refers to the percentage of COD removed by an acidification treatment ahead of AD

2.2. Extraction of extracellular polymeric substances

Solubilization of EPS from the extracellular matrix involved applying an alkaline extraction protocol to lyophilized granules, following the method described by Pinel et al. (Pinel et al., 2020). In brief, lyophilized samples were added to 0.1 M NaOH solution, in a 10 mg/mL concentration, and stirred vigorously for 30 minutes at 80 ^{o}C. The mixture was cooled to 4 ^{o}C and centrifuged at 3300 x g for 30 minutes. The supernatant was collected and dialyzed with a 3.5 kDa cut-off dialysis bag against demi-water overnight at room temperature and, subsequently, lyophilized.

2.3. General EPS characterization

Colorimetric carbohydrate and protein quantification in the granular EPS was performed as described by Dubois et al. (Dubois et al., 1951). In brief, lyophilized extracts obtained as described above were resolubilized in 0.01 M NaOH to a final concentration of 0.5 mg/mL. The protein content was measured in triplicate using the Pierce™ BCA protein assay kit (Thermo Scientific, Rockford, IL) according to the manufacturer's instruction. Bovine serum albumin was used as a standard. Carbohydrate content was measured in triplicate with phenol-sulfuric acid method (Dubois et al., 1951), using glucose as a standard.

Fourier-transform infrared spectroscopy (FT-IR) was used to measure the absorbance of functional groups in the lyophilized EPS extracts in triplicate with a Spectrum 100 spectrometer (PerkinElmer, Shelton, CT). The absorbance of the EPS extracts was recorded with FT-IR in attenuated total reflectance mode over a wavenumber range of 600 - 4000 cm^{-1} with 16 accumulations and a 4 cm^{-1} resolution. MATLAB was used for spectral data processing and consisted of baseline correction and feature scaling.

2.4. Hydrolysis and sugar monomer quantification

Composition of sugar monomers in the EPS was determined with high-performance anion exchange with pulsed amperometric detection (HPAE-PAD). Sugar monomers were released from polysaccharides through hydrolysis following the procedure described by Felz et al. (Felz et al., 2019). In brief, samples were hydrolyzed using 1 M hydrochloric acid at a concentration of 10 grams of sample per liter of hydrochloric acid. Hydrolysis was performed at 105 °C for 8 hours in a heating block. Samples were centrifuged at 13,300 x g for 5 min and were neutralized with equal volume of 1M sodium hydroxide. The samples were diluted 1:5 with ultrapure water and filtered through a 0.22 μm PVDF filter. Quantification of the sugar monomers was performed by calibration of sugar monomers in a range of 0.025 g/L to 0.2 g/L. The sugar monomer standards used were: galacturonic acid, glucuronic acid, rhamnose, glucose, glucosamine, mannose, galactose and fucose. The detection of sugar monomers in the standards and hydrolyzed EPS samples in triplicates was performed with a Dionex ICS-5000+, equipped with a CarboPac PA20 column and an Aminotrap pre-column as mentioned in Felz et al. (Felz et al., 2019).

2.5. Glycosyl composition analysis by TMS method

Lyophilized EPS samples were sent to the Complex Carbohydrate Research Center for an additional measurement of the sugar monomer composition in EPS (Athens, GA, USA). The method is optimized for measuring sugar monomers in lipo-polysaccharides. The analysis was performed by combined gas chromatography-mass spectrometry (GC-MS) of the *O*-trimethylsilyl (TMS) methyl glycoside derivatives produced from the sample by acidic methanolysis. These procedures were carried out as previously described (Santander et al., 2013). In brief, 300 mg of lyophillized EPS samples, and 20 mg of inositiol as internal standard, were added to separate tubes. Samples were hydrolyzed in 1M HCl in methanol at 80 °C for 16 hours. This was followed by re-N-acetylation of amino-sugars with pyridine and acetic anhydride in methanol. Subsequently, samples were per-o-trimethylsilyated with Tri-Sil reagent at 80 °C for 30 minutes. GC-MS analysis of the TMS methyl glycosides was performed on an AT 7890A GC interfaced to a 5975B MSD, using an EC-1 fused silica capillary column (30 m × 0.25 mm ID).

2.6. DNA extraction and sequencing

Total genomic DNA was extracted from refrigerated sludge samples (+4°C) using DNeasy PowerSoil Pro kit (Qiagen, Germany), following the manufacturer's protocol. Library preparation and sequencing with Illumina NovaSeq 6000 PE150 was done at Novogene (Cambridge, UK), yielding on average 20Gb of raw data per sample. Raw reads were deposited to the European Nucleotide Archive (ENA) under the study accession number ERP148022. Paired-end reads were trimmed for the adapter and quality sequences and assembled with MEGAHIT v1.2.9 (Li et al., 2015). Trimmed reads were mapped to the assembled contigs with bowtie2 in BBMap suite v38.84 (Bushnell, 2015) and used for differential coverage binning with MetaBAT2 (Kang et al., 2015). The quality of the resulting metagenome assembled genomes (MAGs) were checked with CheckM (Parks et al., 2015) using the lineage-specific workflow. Taxonomic affiliation of the MAGs was done using GTDB-Tk 1.7.0 (Chaumeil et al., 2020).

2.7. Annotation of the assembled metagenomes

Annotation of MAGs was done with Prokka 1.14.6 (Seemann, 2014) with the kingdom-specific annotation mode (Bacteria or Archaea), based on the phylogenetic affiliation provided by GTDB-Tk. For the analysis of the biosynthetic potential, only the MAGs passing these quality criteria were used: presence of SSU rRNA (small subunit ribosomal ribonucleic acid gene(s)) with length of more than 1000bp), presence of essential marker genes (Parks et al., 2015), bin completeness of more than 70% with contamination of less or equal to 5%. MAGs passing these criteria were called "qualified MAGs" for the correspondent samples. The relative abundance of each MAG was estimated by calculating the proportion of mapped reads to each MAG relative to total amount of reads.

Additional re-annotation of the metagenomes was done through the MG-RAST (Meyer et al., 2008) to place genes to the distinct metabolic sub-systems. Amino acid

sequences from the translated sequenced genes were classified as belonging to the specific subsystems with the E-value threshold of 1e-5 and sequence identity of more than 60%. Quality-filtered reference MAGs were additionally re-annotated in RAST (Overbeek et al., 2014). Automatic annotations were manually curated with the assistance of KEGG (Kyoto Encyclopedia of Genes and Genomes, Release 103.0, July 1, 2022 (Kanehisa et al., 2014)).

To identify carbohydrate-active enzymes that might be involved in the biosynthesis of EPS, translated protein sequences of the qualified MAGs were mapped to the CAZy database (Drula et al., 2022) with dbCAN-fam-HMMs.txt (dbCAN release 4.0) with criteria of E-value of less than 1e-5 and sequence coverage of more than 0.35 (Xia et al., 2013; Yin et al., 2012). For each sample unique CAZy instances belonging to one of the groups (glycoside hydrolases, glycosyl transferase, polysaccharide lyases, carbohydrate esterases, carbohydrate-binding modules) were calculated and averaged for all reference MAGs from that sample. Genes and proteins for the biosynthesis of the specific sugar nucleotides were selected by querying the MetaCyc database (Caspi et al., 2020) with the search terms like "rhamnose biosynthesis", "UDP-sugar biosynthesis", "alginate biosynthesis" and selecting the relevant Bacterial or Archaeal pathways. Amino acid sequences of the selected genes (Supplementary Spreadsheet 1) were then used to query the MAGs with BlastP using the E-value of less than or equal to 1e-2 as a cutoff (McGinnis and Madden, 2004).

3. Results

In this study we characterized chemical and microbiological composition of three different granular anaerobic sludges that originated from two types of organic industrial wastewater: paper recycle/kraft, and brewery. The two IC bioreactors treating paper recycle/kraft wastewater differed by the type of generated sludge: fluffy granular (Papermill A) and smooth granular (Papermill B); while the UASB reactor treating brewery wastewater had smooth granular sludge.

3.1. General characterization of the granular sludge extracellular matrix

Conventional colorimetric carbohydrate and protein quantification (see section 2.3) was used to get a general distribution of two of the major constituents of EPS in the sludges: polysaccharides and proteins. Proteins were predominant in the EPS of all sludges, but especially in the papermill granules. The total carbohydrate content was 9.9 ± 1.5, 9.5 ± 0.3 and 6.5 ± 0.8 % (glucose equivalent) of total EPS in granules from Papermill A, Papermill B and Brewery EPS, respectively. The PN/PS ratio was 9.4 ± 1.4, 9.8 ± 0.9, and 8.8 ± 1.4 g BSA-eq/g glucose-eq for EPS in granules from Papermill A, Papermill B and Brewery.

The quantification of the total carbohydrates fraction of EPS obtained with the colorimetric methods was validated by measuring the relative abundance in functional groups with FT-IR (Figure 1). An overview of the assigned wavenumbers corresponding to the functional groups can be found in Table S1. Bands corresponding to protein absorbance regions were dominant in the spectrum (bands 1, 2, 6, 7, 9 and 11) and were most pronounced in the papermill samples. The measured absorbance height in the polysaccharide region (band 12) was 3% higher in the spectrum of the Brewery sludge compared to both Papermill sludge samples. This showed that based on the absorbance height in the FT-IR spectra, the overall EPS composition in the granules is similar, but their polysaccharide content slightly varies.

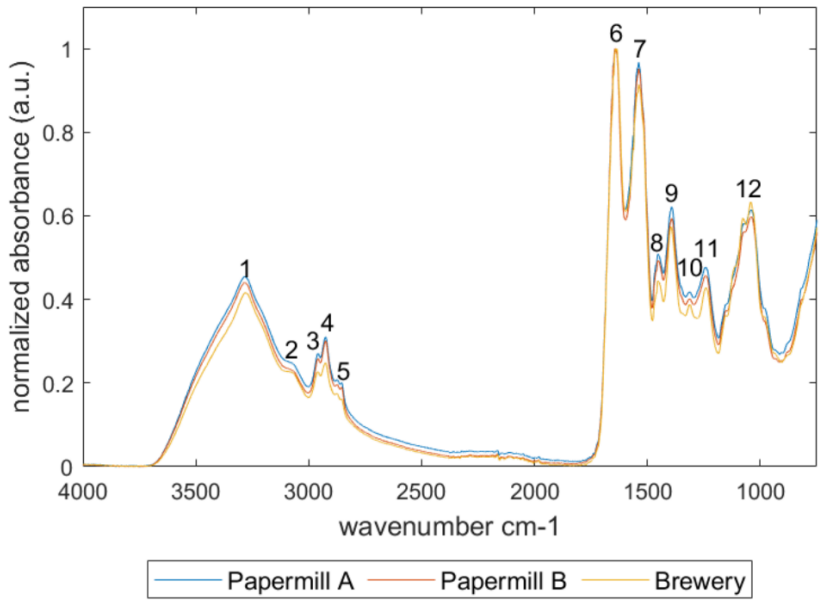


Figure 1. Normalized absorbance FT-IR spectra of EPS extracted from Papermill A, Papermill B and Brewery granular sludge.

3.2. Characterization of glycan composition in the extracellular matrix

Total sugar content in the EPS varied slightly between the sludge samples (Figures S2, S3, Table S2): 9.0 ± 0.8, 12.2 ± 1.1 and 10.1 ± 1.8 % of total EPS, for Papermill A, Papermill B and Brewery samples, respectively. The sugar monomer composition of the two Papermill samples was similar, with differences in the galactosamine, glucosamine and galacturonic acid content (Figure 2). Compared to Papermill samples, the Brewery sample had lower glucose and mannose content, but higher content of glucuronic acid.

To get additional information on sugar monomers present in the sludge EPS, TMS-derivatization analysis was performed (Supplemental Figure S4). Since this method is typically used for lipopolysaccharide analysis, differences between the sugar monomers composition obtained by TMS-derivatization and HPAE-PAD method were expected. The hydrolysis of the polysaccharides for TMS-derivatization was performed under less harsh conditions, which was reflected in the total sugar content (4.0 % and 1.9 %). Compared to the HPAE-PAD data (Figure 2), practically no uronic acids were found with TMS-derivatization. This could indicate that the hydrolysis method used for the TMS-derivatization might not be strong enough for the hydrolysis of all glycans present in the EPS.

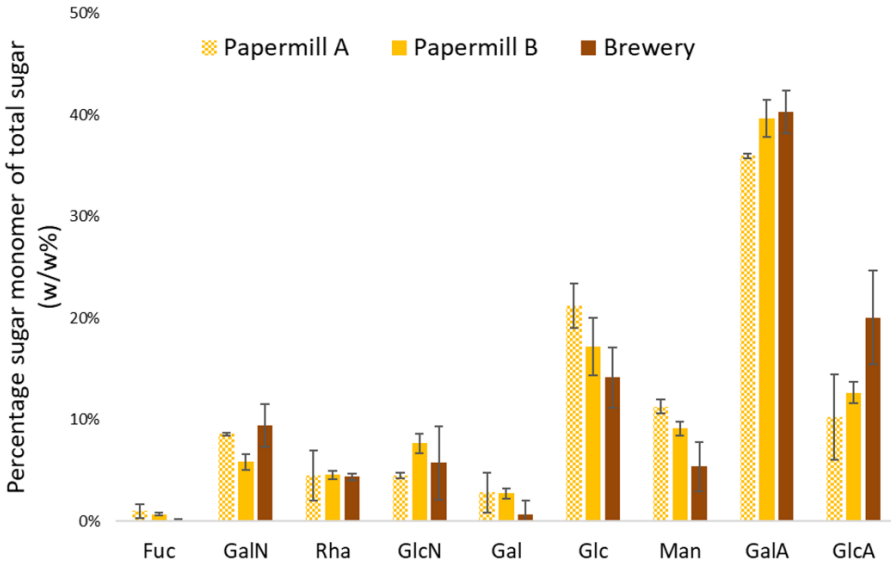


Figure 2. Sugar monomer composition of EPS measured with HPAE-PAD chromatography. The sugar monomers are displayed as a percentage of the total sugar concentration found in the EPS. Abbreviations are according to symbol nomenclature for glycans (Varki et al., 2015). Fu stands for fucose, GalN for galactosamine, Rha for rhamnose, GlcN for glucosamine, Gal for galactose, Glc for glucose, Man for mannose, GalA for galacturonic acid and GlcA for glucuronic acid. See Supplemental Table S3 for the raw data. The error bars represent the standard deviation.

3.3. Metagenome assembly and taxonomic classification

Analysis of the sludges metagenome allowed to establish the link between EPS composition and microbial community. Extracted DNA from each granular sludge sample was sequenced to the same depth, generating 69.9 Gb of raw data in total, with 140-160 million raw reads per sample. Quality scores were above 91% for each of the samples in Q30.

The highest number of MAGs were retrieved from the Papermill A sample (113), followed by Papermill B (117) and Brewery (106). MAGs from Papermill A had a higher mean completeness (78%), while Papermill B had the lowest (42%) (Figure S5, A). All samples shared low (less than 5%) contamination of the MAGs with the genetic material of the different phylogenetic placement (Figure S5, B).

Taxonomic affiliation of the 16S rRNA gene sequences in the metagenomes showed that samples from IC reactors (Papermill A and B) were more similar than the UASB Brewery sample (Figure 3). Brewery-retrieved metagenomes had more 16S rRNA gene sequences belonging to the *Proteobacteria* phylum and less belonging to the *Firmicutes*, *Chloroflexi* and *Actinobacteria* phyla. Interestingly, the number of 16S rRNA

gene sequences belonging to the *Euryarchaeota* phylum was twice as high in the Brewery granular sludge, then in either Papermill samples.

Normalized abundance and phylogenetic placement of all the retrieved MAGs from the three anaerobic sludge samples (Figure S6) demonstrates a diverse range of dominant taxa across the three samples, regardless of the shared bioreactor specification (Papermill A and B samples), or similar granulation state of the sludge (Papermill B and Brewery sample, smooth granules). The shared and distinct phylogenetic groups were visualized by constructing a Venn diagram (Figure S7). Notably, 10 MAGs belonging to the *Anaerolineaceae* family were shared among all three samples, as well as 2 MAGs from *Syntrophobacteraceae* and one belonging to the *Desulfovibrionaceae* family (Supplementary Spreadsheet 1). Apart from these, distinct MAGs were shared across the three samples belonging to the classes of *Clostridia*, *Clostridia*_A, *Actinomycetia*, *Syntrophorhabdia*, *Verrucomicrobiae*, *Syntrophia*, *Thermoleophilia*, *Spirochaetia*, *Syntrophobacteria*, *Phycisphaerae*, *Bacteroidia* and *Desulfovibrionia*. Among MAGs classified as Archaea, those belonging to the classes *Bathyarchaeia*, *Methanosarcinia* and *Methanobacteria* were found in all three samples, while Brewery sample had distinct MAGs belonging to the *Methanomicrobia* and *Thermococci* classes, and Papermill A sample had also MAGs belonging to *Thermoplasmata* class (specifically, to *Methanomassiliicoccus* genus). MAG classified as *UBA233* family within *Bathyarchaeia* class was found in all three anaerobic sludge samples.

Across all three samples, MAGs belonging to the *Anaerolineaceae* family had the highest bacterial coverage, with an average of 100-400 reads assigned per MAG and representing 9-16% of all mapped reads. For Archaea, MAGs belonging to the genus *Methanobacterium* (400 – 900 reads per MAG) had the highest coverage.

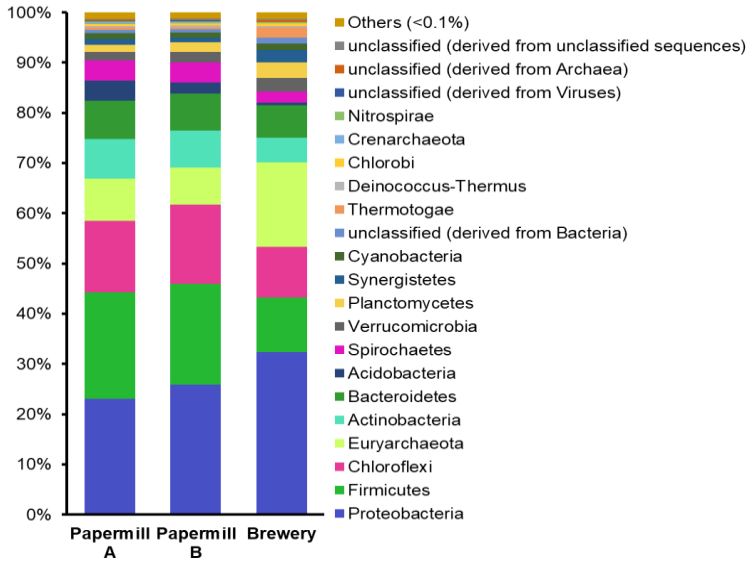


Figure 3. Taxonomic classification of the metagenome sequences from the three anaerobic sludge samples on the phylum level.

3.4. Analysis of the pangenome across the three sludge samples

Gene-centric analysis was used to identify shared and distinct gene functions within the three metagenomes. Addressing the glycan and EPS focus of this study, we investigated the distribution of the genes associated with the metabolism of polysaccharides (Figure 4). Four sub-systems were identified with MG-RAST and KEGG (see Methods) to harbor the relevant genes, within the overarching "capsular polysaccharides biosynthesis and assembly": (1) Exopolysaccharide biosynthesis, (2) Rhamnose containing glycans, (3) dTDP-rhamnose synthesis and (4) Alginate metabolism. Brewery MAGs had more genes classified within "capsular polysaccharides biosynthesis and assembly" (1913 genes) than either of the Papermill samples (1679 (A) and 1799 (B) genes). High number of genes with the general glycosyltransferase (GT) functions (Figure S8) explains a diversity of the sugar monomers identified in the HPAE-PAD analysis of the sludges EPS (Figure 2). Meanwhile, biosynthesis of monosaccharides, such as rhamnose and its precursor dTDP-rhamnose, is almost evenly distributed across the three sludges (Figure S9, S10), matching the uniform distribution of rhamnose sugar monomer content in the three samples (Figure 2).

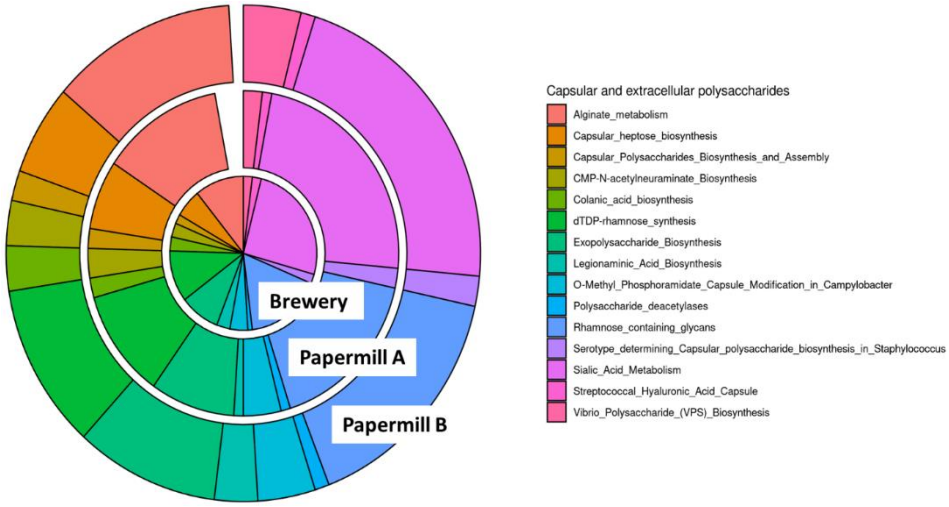


Figure 4. Genes identified in the three sludge samples related to the metabolism of capsular and extracellular polysaccharides.

Interestingly, the heteropolysaccharide biosynthesis potential of the MAGs revolved around alginate metabolism cluster, which was 30% more abundant in smooth sludge (Papermill B and Brewery) compared to the floccular (Papermill A). However, only Papermill B sample had the full operon for the synthesis of alginate (Figure 5).

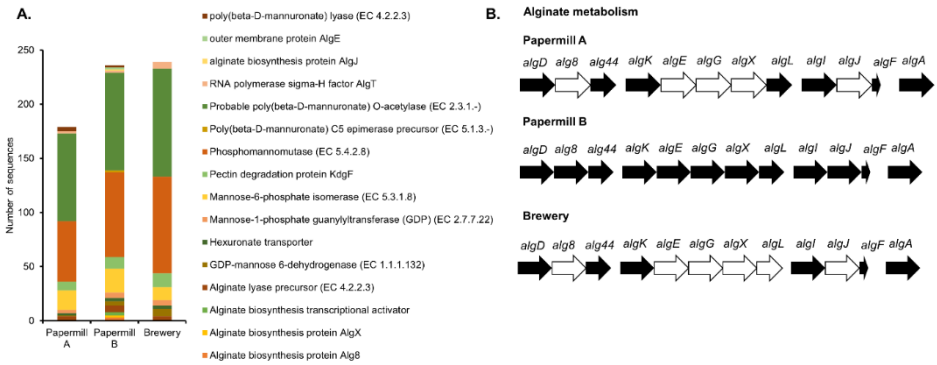


Figure 5. Genes identified in the three sludge samples related to the alginate metabolism (A) and presence/absence of these genes within the alginate operon (B).

3.5. Carbohydrate Active Enzymes across the sludge MAGs

Since the majority of the EPS-associated genes in the three sludges were glycosyltransferases of diverse functions (Figure S7), we looked closer into their distribution among MAGs to pinpoint potential EPS-producing microorganisms. GTs represent one of the classes in the carbohydrate-active enzymes (CAZymes) in CAZy database (http://www.cazy.org/, (Drula et al., 2022)), together with glycoside hydrolases (GH), polysaccharide lyases (PL), carbohydrate esterases (CE) and carbohydrate-binding modules (CBM). Enzymes within GT group catalyze transfer of sugar moieties from donor to acceptor molecules, resulting in the formation of glycosidic bonds in the polysaccharides. Gene clusters associated with the synthesis of the known EPS molecules (ex. xanthan, gellan, succinoglycan, alginate) always have one or multiple GT-encoding genes (Sun and Zhang, 2021).

Qualified MAGs, with more than 70% completeness and less than 5% contamination, had a diverse representation of CAZymes across all classes in all three sludges (Figure S11). Qualified MAGs in Papermill A with the greatest number of CAZymes belonged to the bacterial classes *Phycisphaerae* (MAG36), *UBA5829* (MAG122), *Bacteroidia* (MAG124) and *UB6911*(MAG125). While glycoside hydrolases were the predominant CAZymes in the floccular sample Papermill A, glycosyltransferases were the most common in Papermill B and Brewery MAGs. Among the archaeal qualified MAGs, Brewery sample had the most CAZymes encoded in MAG20 (*Methanomicrobia*), MAG116 (*Methanobacteria*) and MAG123 (*Methanomicrobia*). Interestingly, MAG116 had almost twice as many CAZymes as any other archaeal MAGs from the three samples, with most in the glycosyltransferase family.

Most abundant MAGs (with the coverage above 100 reads) in Papermill A and B samples, *Anaerolineae* MAG22, 29, and 89 had slightly lower number of CAZymes than qualified MAGs from the same samples (Figure 6). Glycosyltransferases were represented by the GT2 and GT4 groups, that include enzymes involved in the biosynthesis of lipopolysaccharides and osmoprotectants

(mannosylglucosylglycerate synthase). Presence of glycogen phosphorylase in MAG22 hints on the potential of this microorganism to synthesize an intracellular starch-like polysaccharide, glycogen. MAG22 also encoded cellobiose phosphorylase, supporting earlier claims that members of the *Anaerolineae* class can ferment carbohydrates, such as cellobiose and cellulose. Function of the CAZymes identified in the MAG29 and 89 were similar to the MAG22, but with more glycosyltransferases involved in the transport of mannose and galacturonic acid, as part of the lipid and lipopolysaccharide biosynthesis pathways (polyprenol monophosphomannose synthases and dodecaprenyl-phosphate galacturonate synthases). MAG89 also had diverse N-acetylglucosaminyltransferases, which might be used for the biosynthesis of poly-beta-1,6-N-acetyl-D-glucosamine, a polysaccharide intercellular adhesin (Arciola et al., 2015).

Most abundant MAGs in the Brewery sample also had the highest number of CAZymes: *Methanosarcinia* MAG108 and *Methanobacteria* MAG122 (Figure 6, Figure S10). 60% of the GTs in MAG108 and MAG122 belonged to the GT41 group, that includes β-N-acetylglucosaminyltransferases, N-β-glucosyltransferases and O-α-L-fucosyltransferases.

It is important to also note that 40-50% of the identified glycosyltransferases in all the MAGs described here did not have exact functional roles assigned (aka. putative glycosyltransferases).

		GT	GH	PL	CE	CBM	Completeness
Papermill A	Methanobacteria bin 17	34	2	0	16	0	●
	Anaerolineae bin 22	38	13	4	12	10	◑
	Anaerolineae bin 29	60	32	2	18	12	◑
	Bacteroidia bin 40	23	67	3	22	11	◕
	Actinomycetia bin 62	18	16	0	20	5	●
	Methanobacteria bin 63	23	2	0	5	0	●
	Clostridia bin 102	19	68	2	23	13	●
	Methanosarcinia bin 106	41	5	2	10	3	◑
Papermill B	Bacteroidia bin 12	21	47	2	13	9	◑
	Clostridia bin 23	17	56	2	12	8	◕
	Anaerolineae bin 89	48	27	4	14	12	◑
	Anaerolineae bin 100	26	18	4	10	9	◑
	Actinomycetia bin 107	16	14	0	16	5	●
Brewery	Methanobacteria bin 55	22	5	0	8	2	◑
	Desulfuromonadia bin 64	21	3	0	0	2	◑
	Methanosarcinia bin 108	34	6	4	4	2	◕
	Methanobacteria bin 122	30	2	3	14	0	◕

Figure 6. Number of CAZymes from the most abundant MAGs across all three anaerobic sludge samples. GT – Glycosyl Transferases, GH – Glycoside Hydrolases, PL – Polysaccharide Lyases, CE – Carbohydrate Esterases, CBM – Carbohydrate-Binding Modules.

3.6. Biosynthesis of sugar monomers by the abundant MAGs

To match the HPAE-PAD identified sugar monomers with the MAGs biosynthetic capabilities, we analyzed the metagenomes for the presence of genes for the biosynthesis of correspondent precursor nucleotide sugars. Nucleotide sugars, as activated monosaccharides, serve as important intermediates to the production of polysaccharides (Franklin et al., 2011). As a result, we found complete operons for the biosynthesis of UDP-α-D-galacturonate (UDP-GalA), UDP-N-acetyl-D-galactosamine (UDP-GalNAc) and dTDP-β-L-rhamnose (dTDP-l-Rha) in most of the abundant MAGs (Figure 7).

		UDP-GalA	UDP-GalNAc	GDP-Man	GDP-Fuc	UDP-L-FucNAm	dTDP-l-Rha	GDP-Rha	NulO-Sia
Papermill A	Methanobacteria bin 17	5	12	6	3	4	11	3	0
	Anaerolineae bin 22	12	12	1	6	8	16	7	1
	Anaerolineae bin 29	13	17	1	6	7	15	10	5
	Bacteroidia bin 40	6	8	4	3	5	6	2	1
	Actinomycetia bin 62	2	8	4	1	3	4	3	0
	Methanobacteria bin 63	2	11	7	1	4	9	2	0
	Clostridia bin 102	8	18	6	5	1	13	6	1
	Methanosarcinia bin 106	10	7	3	6	7	12	7	0
Papermill B	Bacteroidia bin 12	3	4	3	3	3	4	3	0
	Clostridia bin 23	3	11	3	4	2	4	4	1
	Anaerolineae bin 89	12	17	3	7	9	18	13	2
	Anaerolineae bin 100	9	15	1	7	6	12	8	0
	Actinomycetia bin 107	3	8	3	2	1	3	4	0
Brewery	Methanobacteria bin 55	5	12	4	3	2	8	2	1
	Desulfuromonadia bin 64	3	5	1	2	2	7	2	0
	Methanosarcinia bin 108	3	7	2	2	2	6	5	0
	Methanobacteria bin 122	4	11	5	3	4	9	3	0

Figure 7. Number of glycan-associated genes found across the abundant MAGs. UDP-GalA – biosynthesis of UDP-α-D-galacturonate; UDP-GalNAc – biosynthesis of UDP-N-acetyl-D-galactosamine; GDP-Man – biosynthesis of GDP-mannose; GDP-Fuc – biosynthesis of GDP-fucose; UDP-L-FucNAm – biosynthesis of UDP-N-acetyl-β-L-fucosamine; dTDP-l-Rha – biosynthesis of dTDP-β-L-rhamnose; GDP-Rha – biosynthesis of GDP-D-rhamnose; NulO-Sia – biosynthesis of NulO sialic acid.

Abundant MAGs in the Papermill A and B sludges classified as *Anaerolineae*, all had a high number of nucleotide sugar biosynthesis genes, matching the GT2 and GT4 classes of glycosyltransferases predicted with the CAZy search (Figure 6). These MAGs also had genes for the biosynthesis of nonulosonic acids (NulO-Sia). Abundant MAGs of methanogenic archaea also had a potential to synthesize UDP-GalNAc and dTDP-l-Rha, while biosynthetic potential for GDP-Rha was more characteristic to the bacterial MAGs.

Since in the pangenome analysis we identified complete operons for the biosynthesis of alginate (in Papermill B sludge (Figure 5), we checked whether individual abundant MAGs had the capacity to synthesize this polysaccharide. However, we only identified a few genes in the alginate biosynthesis pathway such as GDP-mannose 6-dehydrogenase (*algD*) and mannuronan C5-epimerase (*algG*).

4. Discussion

Extracellular polymeric substances produced by anaerobic microorganisms remain vastly unknown and poorly characterized. Despite a widespread industrial application of microbial biotechnologies relying on EPS-based aggregate formation, chemical and biological make-up of these micro-factories are rarely linked in a comprehensive way. In this study we investigated the chemical composition of EPS from three anaerobic granular sludges, coming from papermill and brewery wastewater treatment facilities, and used the chemical EPS composition to guide the assessment of EPS biosynthetic potential in the sludge microbiomes.

4.1. Importance of glycan analysis in protein dominated EPS

The PN/PS ratio of 8.8-9.8 (g BSA-eq/g glucose-eq) demonstrates that EPS from papermill and brewery granular sludges was protein-rich. High protein content is typical for slow-growing microbial communities and similar PN/PS ratios were found in anammox granular sludges (~10 g BSA-eq/g glucose-eq) (Lotti et al., 2019). Higher protein concentrations were linked to a higher hydrophobicity and granule formation capacity (Santschi et al., 2020). Liu and Fang suggested a positive correlation between higher Archaeal abundance and PN/PS ratio, since methanogenic sludge had 10x PN/PS ratio compared to the acidogenic one (Liu and Fang, 2002). Although proteins are often the dominant part of studied EPS, in the extracellular matrix they are often linked with polysaccharides that can alter their structural properties, offer protection from digestion and aid in osmotic regulations (Varki et al., 2022). Since polysaccharides (glycans) can constitute 40-95% of the overall extracellular matrix (More et al., 2014), their diversity in the granular sludge may play a central role in defining the granules' activity and microbial interconnections.

4.2. Glycan composition of the granular sludge

In previously studied non-granular microbial biofilms, both bacteria and archaea produced a wide array of unique glycans in the process of microbial glycosylation that bears distinct steps from the eukaryotic ones (Dell et al., 2010). These diverse microbial glycans can be commonly found attached to lipids, proteins or loosely floating in the extracellular matrix; while the most dominant ones are associated with the microbial cell-surface and include mono- and heteropolysaccharides (Barrett and Dube, 2023; Messner et al., 2010). The complexity of the microbial communities and the added diversity of glycans produced by the microbial community make isolation and identification of single glycans challenging. Therefore, to be able to characterize the full spectrum of the microbially-produced glycans in the tested here granular sludges, we hydrolyzed the total extracted granular EPS to identify the monosaccharides composition of the microbe-associated glycans. The full hydrolysis applied to the samples in this study allowed to overcome the inconsistent reports addressing EPS composition in granular sludge from the papermill and brewery wastewater treating bioreactors (Gonzalez-Gil et al., 2015; Liu and Fang, 2002). In support of the previously reported high molecular weight of the granular sludge EPS (Gonzalez-Gil et al., 2015), we indeed observed that 16.9 – 27.4 % of the total EPS

from papermill and brewery granular sludge studied here had a molecular weight of over 5500 kDa (Chen et al., 2023).

Distribution of monosaccharides in papermill and brewery sludges (Figure 2) partially matches the monosaccharide diversity reported for other anaerobic granular sludges. For example, Veiga et al. (Veiga et al., 1997) showed that the monosaccharide composition of in anaerobic granules composed of syntrophic acetogenic bacteria and methanogenic archaea contained glucose (19.3%), rhamnose (15.0%), mannose (12.9%), fucose (10.7%), galactose (10.1%), glucosamine (8.6%) and galactosamine (6.4 %). While papermill and brewery sludges studied here were more microbiologically diverse (Figure 3), the most dominant identified sugar monomers were negatively charged uronic acids (46-60%), and not glucose (Figure 2). Presence of other negatively charged glycans, was previously reported for the similar granular sludges to the ones studied here (papermill and brewery) (Gonzalez-Gil et al., 2015). Using ^{1}H NMR and MALDI-TOF analyses, authors reported presence of spectral signatures of mannuronic acid, which is closely related to alginate (Gonzalez-Gil et al., 2015). While the sugar monomer composition demonstrated here shows presence of glucuronic and galacturonic acids (Figure 2), their molecular weight is similar to the previously found mannuronic acid (194 g/mol), making the results comparable to the previous reports. Also consistent with previous findings (Liu and Fang, 2002) is the presence of higher concentrations of uronic acids in the floccular sludge (Papermill A) compared to the smooth granular sludges (Papermill B and Brewery).

4.3. Microbial glycan biosynthetic potential

Analysis of the sugar monomers composition of the granular sludge with HPAE-PAD was then used to guide the search for the glycans biosynthetic genes in the sludges microbiomes. By sequencing the metagenomes of the three anaerobic granular sludges we were able to look closely into the glycan biosynthetic potential of the assembled high quality abundant MAGs (Figure 6, 7). The combined analysis of the MAGs-encoded CAZymes and biosynthesis genes for the glycans precursors, nucleotide-sugars, demonstrated matching results for the two search strategies. Similarly to the earlier reported metagenome analysis of papermill and brewery anaerobic granular sludges (Gonzalez-Gil et al., 2015), we identified more alginate-associated biosynthesis genes in the Papermill metagenome than in the Brewery. However, we were not able to pinpoint the exact MAGs that contained the full alginate biosynthetic gene cluster. A number of abundant MAGs in all three sludges shared instead a common potential for the biosynthesis of galacturonate, N-acetyl-D-galactosamine and rhamnose (Figure 7). Based on HPAE-PAD analysis, these sugars were also the most abundant monosaccharides contributing to the glycan part of the three anaerobic sludges (Figure 2). Therefore, it is plausible to suggest that the MAGs with the identified genes for the biosynthesis of these sugar nucleotides are involved in the formation of sludge extracellular matrix.

Strikingly, the MAGs with the highest number of glycan and nucleotide sugars biosynthetic capabilities were also the most abundant across the three samples and belonged to the *Anaerolineae* and *Methanobacteria* classes. Their high abundance is

supported by other studies on the microbial composition of various mesophilic anaerobic sludges, where *Methanobacteria* were the most common and abundant hydrogenotrophic methanogens (Griffin et al., 1998; McHugh et al., 2003; Song et al., 2010), while *Anaerolineae* contributed to the 45% of the overall anaerobic microbial sludge population in the fed-batch, UASB and IC reactors treating lipid, alkane-rich or papermill wastewater, as well as waste activated sludge (Liang et al., 2015; McIlroy et al., 2017; Rivière et al., 2009). Some members of *Anaerolineae* were even found in the beneficial trophic relationships with the hydrogenotrophic methanogens from the genus *Methanospirillum* (Sun et al., 2016). Currently known representatives of *Anaerolineaceae* family are anaerobic nonmotile slow growing acetogens, with chemoorganoheterotrophic fermentative metabolism and multicellular filamentous morphology (Yamada and Sekiguchi, 2018). Although common in the dispersed anaerobic sludges, morphology of these representatives of *Chloroflexota* phylum can be contributing to the anaerobic sludge granulation (Sekiguchi et al., 2001; Yamada et al., 2005), while their ability to produce acetate may contribute to the beneficial association with the acetoclastic methanogens to further improve the strength of the anaerobic granular structures. What is more, Zhu et al. (Zhu et al., 2017) proposed that the filamentous structure of the *Anaerolineaceae* cells may be essential to maintain the granular stability during the temperature and pH fluctuations inside the UASB reactors. High abundance of the *Anaerolineaceae* (up to 30%) and *Methanosaeta*-classified microorganisms (up to 60%) in the sludge treating starch-rich wastewater ensured stability of the bioreactor operation even after the bioreactor pre-acidification (Wu et al., 2021). High number of CAZymes and genes for the production of nucleotide-sugars (glycans precursors) in the abundant *Anaerolineae* and *Methanobacteria* identified in this study suggest a strong link between these microorganisms and the observed glycan profile of the tested anaerobic granular sludges from IC and UASB reactors treating papermill and brewery wastewaters. However, further experimental evidence for the glycan producing activity of these microorganisms (*Anaerolineae* and *Methanobacteria*) is still needed.

5. Conclusions

In this study, by combining chemical characterization of glycans in anaerobic granular sludge with metagenome analysis, it was possible to find genes that might be involved in the biosynthesis of these glycans in anaerobic granular sludge microbial communities. The presence of these genes allowed for the identification of microorganisms that potentially take part in the production and secretion of glycans in anaerobic granular sludges from industrial wastewater treating facilities. The approach presented here shows how the chemical analysis of the sugar monomers in EPS can aid in identifying the putative EPS biosynthesis genes in the sludge MAGs. Functions of the identified carbohydrate-active enzymes encoded in the sludges MAGs were also in line with the presence of specific biosynthesis genes for the glycans precursors, nucleotide-sugars, which were identified as prevalent components of the sludge EPS. Gaining a deeper understanding of the glycan biosynthetic potential of the microorganisms in anaerobic granular sludge is essential

for future endeavors to optimize biological production of specific EPS and improve resource recovery from wastewater.

Acknowledgements
This work was supported by the SIAM Gravitation Grant 024.002.002, The Netherlands Organization for Scientific Research as well as the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Chemical Sciences, Geosciences and Biosciences Division, under award #DE-SC0015662. AD was funded by Dutch Research Council (NWO) grant number OCENW.XS21.4.067, and Wageningen Graduate Schools Postdoc Talent Programme.

References

Arciola, C.R., Campoccia, D., Ravaioli, S., Montanaro, L., 2015. Polysaccharide intercellular adhesin in biofilm: structural and regulatory aspects. Front Cell Infect Microbiol 5. https://doi.org/10.3389/fcimb.2015.00007

Barrett, K., Dube, D.H., 2023. Chemical Tools to Study Bacterial Glycans: A Tale from Discovery of Glycoproteins to Disruption of Their Function. Isr J Chem 63. https://doi.org/10.1002/ijch.202200050

Boleij, M., Kleikamp, H., Pabst, M., Neu, T.R., van Loosdrecht, M.C.M., Lin, Y., 2020. Decorating the Anammox House: Sialic Acids and Sulfated Glycosaminoglycans in the Extracellular Polymeric Substances of Anammox Granular Sludge. Environ Sci Technol 54, 5218–5226. https://doi.org/10.1021/acs.est.9b07207

Bushnell, B., 2015. BBMap [WWW Document]. URL sourceforge.net/projects/bbmap/

Caspi, R., Billington, R., Keseler, I.M., Kothari, A., Krummenacker, M., Midford, P.E., Ong, W.K., Paley, S., Subhraveti, P., Karp, P.D., 2020. The MetaCyc database of metabolic pathways and enzymes - a 2019 update. Nucleic Acids Res 48, D445–D453. https://doi.org/10.1093/nar/gkz862

Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., Parks, D.H., 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36, 1925–1927. https://doi.org/10.1093/bioinformatics/btz848

Chen, L.M., de Bruin, S., Pronk, M., Sousa, D.Z., van Loosdrecht, M.C.M., Lin, Y., 2023. Sialylation and Sulfation of Anionic Glycoconjugates Are Common in the Extracellular Polymeric Substances of Both Aerobic and Anaerobic Granular Sludges. Environ Sci Technol. https://doi.org/10.1021/acs.est.2c09586

Chirman, D., Pleshko, N., 2021. Characterization of bacterial biofilm infections with Fourier transform infrared spectroscopy: a review. Appl Spectrosc Rev 56, 673–701. https://doi.org/10.1080/05704928.2020.1864392

de Bruin, S., Vasquez-Cardenas, D., Sarbu, S.M., Meysman, F.J.R., Sousa, D.Z., van Loosdrecht, M.C.M., Lin, Y., 2022. Sulfated glycosaminoglycan-like polymers are present in an acidophilic biofilm from a sulfidic cave. Science of The Total Environment 829, 154472. https://doi.org/10.1016/j.scitotenv.2022.154472

Decho, A.W., Gutierrez, T., 2017. Microbial Extracellular Polymeric Substances (EPSs) in Ocean Systems. Front Microbiol 8. https://doi.org/10.3389/fmicb.2017.00922

Dell, A., Galadari, A., Sastre, F., Hitchen, P., 2010. Similarities and Differences in the Glycosylation Mechanisms in Prokaryotes and Eukaryotes. Int J Microbiol 2010, 1–14. https://doi.org/10.1155/2010/148178

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., Terrapon, N., 2022. The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Res 50, D571–D577. https://doi.org/10.1093/nar/gkab1045

Dubois, M., Gilles, K., Hamilton, J.K., Rebers, P.A., Smith, F., 1951. A Colorimetric Method for the Determination of Sugars. Nature 168, 167–167. https://doi.org/10.1038/168167a0

Dueholm, M.K.D., Besteman, M., Zeuner, E.J., Riisgaard-Jensen, M., Nielsen, M.E., Vestergaard, S.Z., Heidelbach, S., Bekker, N.S., Nielsen, P.H., 2023. Genetic potential for exopolysaccharide synthesis in activated sludge bacteria uncovered by genome-resolved metagenomics. Water Res 229, 119485. https://doi.org/10.1016/j.watres.2022.119485

Felz, S., Neu, T.R., van Loosdrecht, M.C.M., Lin, Y., 2020. Aerobic granular sludge contains Hyaluronic acid-like and sulfated glycosaminoglycans-like polymers. Water Res 169, 115291. https://doi.org/10.1016/j.watres.2019.115291

Felz, S., Vermeulen, P., van Loosdrecht, M.C.M., Lin, Y.M., 2019. Chemical characterization methods for the analysis of structural extracellular polymeric substances (EPS). Water Res 157, 201–208. https://doi.org/10.1016/j.watres.2019.03.068

Franklin, M.J., Nivens, D.E., Weadge, J.T., Howell, P.L., 2011. Biosynthesis of the Pseudomonas aeruginosa Extracellular Polysaccharides, Alginate, Pel, and Psl. Front Microbiol 2. https://doi.org/10.3389/fmicb.2011.00167

Gagliano, M.C., Neu, T.R., Kuhlicke, U., Sudmalis, D., Temmink, H., Plugge, C.M., 2018. EPS Glycoconjugate Profiles Shift as Adaptive Response in Anaerobic Microbial Granulation at High Salinity. Front Microbiol 9. https://doi.org/10.3389/fmicb.2018.01423

Gonzalez-Gil, G., Thomas, L., Emwas, A.-H., Lens, P.N.L., Saikaly, P.E., 2015. NMR and MALDI-TOF MS based characterization of exopolysaccharides in anaerobic microbial aggregates from full-scale reactors. Sci Rep 5, 14316. https://doi.org/10.1038/srep14316

Griffin, M.E., McMahon, K.D., Mackie, R.I., Raskin, L., 1998. Methanogenic population dynamics during start-up of anaerobic digesters treating municipal solid waste and biosolids. Biotechnol Bioeng 57, 342–355. https://doi.org/10.1002/(SICI)1097-0290(19980205)57:3<342::AID-BIT11>3.0.CO;2-I

Hulshoff Pol, L.W., de Castro Lopes, S.I., Lettinga, G., Lens, P.N.L., 2004. Anaerobic sludge granulation. Water Res 38, 1376–1389. https://doi.org/10.1016/j.watres.2003.12.002

Jennings, L.K., Storek, K.M., Ledvina, H.E., Coulon, C., Marmont, L.S., Sadovskaya, I., Secor, P.R., Tseng, B.S., Scian, M., Filloux, A.,

Wozniak, D.J., Howell, P.L., Parsek, M.R., 2015. Pel is a cationic exopolysaccharide that cross-links extracellular DNA in the *Pseudomonas aeruginosa* biofilm matrix. Proceedings of the National Academy of Sciences 112, 11353–11358. https://doi.org/10.1073/pnas.1503058112

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M., 2014. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res 42, D199–D205. https://doi.org/10.1093/nar/gkt1076

Kang, D.D., Froula, J., Egan, R., Wang, Z., 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3, e1165. https://doi.org/10.7717/peerj.1165

Kumar, A., Singh, A.K., Bilal, M., Prasad, S., Rameshwari, K.R.T., Chandra, R., 2022. Paper and pulp mill wastewater: characterization, microbial-mediated degradation, and challenges, in: Nanotechnology in Paper and Wood Engineering. Elsevier, pp. 371–387. https://doi.org/10.1016/B978-0-323-85835-9.00011-8

Li, D., Liu, C.-M., Luo, R., Sadakane, K., Lam, T.-W., 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31, 1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Liang, B., Wang, L.-Y., Mbadinga, S.M., Liu, J.-F., Yang, S.-Z., Gu, J.-D., Mu, B.-Z., 2015. Anaerolineaceae and Methanosaeta turned to be the dominant microorganisms in alkanes-dependent methanogenic culture after long-term of incubation. AMB Express 5, 37. https://doi.org/10.1186/s13568-015-0117-4

Limoli, D.H., Jones, C.J., Wozniak, D.J., 2015. Bacterial Extracellular Polysaccharides in Biofilm Formation and Function. Microbiol Spectr 3. https://doi.org/10.1128/microbiolspec.MB-0011-2014

Liu, H., Fang, H.H.P., 2002. Extraction of extracellular polymeric substances (EPS) of sludges. J Biotechnol 95, 249–256. https://doi.org/10.1016/S0168-1656(02)00025-1

Lotti, T., Carretti, E., Berti, D., Martina, M.R., Lubello, C., Malpei, F., 2019. Extraction, recovery and characterization of structural extracellular polymeric substances from anammox granular sludge. J Environ Manage 236, 649–656. https://doi.org/10.1016/j.jenvman.2019.01.054

McGinnis, S., Madden, T.L., 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res 32, W20–W25. https://doi.org/10.1093/nar/gkh435

McHugh, S., Carton, M., Mahony, T., O'Flaherty, V., 2003. Methanogenic population structure in a variety of anaerobic bioreactors. FEMS Microbiol Lett 219, 297–304. https://doi.org/10.1016/S0378-1097(03)00055-7

McIlroy, S.J., Kirkegaard, R.H., Dueholm, M.S., Fernando, E., Karst, S.M., Albertsen, M., Nielsen, P.H., 2017. Culture-Independent Analyses Reveal Novel Anaerolineaceae as Abundant Primary Fermenters in Anaerobic Digesters Treating Waste Activated Sludge. Front Microbiol 8. https://doi.org/10.3389/fmicb.2017.01134

Messner, P., Egelseer, E.M., Sleytr, U.B., Schäffer, C., 2010. Chapter 7 - Bacterial surface layer glycoproteins and "non-classical" secondary cell wall polymers, in: Holst, O., Brennan, P.J., Itzstein, M. von, Moran, A.P. (Eds.), Microbial Glycobiology. Academic Press, San Diego, pp. 109–128. https://doi.org/https://doi.org/10.1016/B978-0-12-374546-0.00007-9

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., Edwards, R., 2008. The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics 9, 386. https://doi.org/10.1186/1471-2105-9-386

More, T.T., Yadav, J.S.S., Yan, S., Tyagi, R.D., Surampalli, R.Y., 2014. Extracellular polymeric substances of bacteria and their potential environmental applications. J Environ Manage 144, 1–25. https://doi.org/10.1016/j.jenvman.2014.05.010

Neu, T., Kuhlicke, U., 2017. Fluorescence Lectin Bar-Coding of Glycoconjugates in the Extracellular Matrix of Biofilm and Bioaggregate Forming Microorganisms. Microorganisms 5, 5. https://doi.org/10.3390/microorganisms5010005

Neu, T.R., Lawrence, J.R., 2010. Extracellular polymeric substances in microbial biofilms, in: Microbial Glycobiology. Elsevier, pp. 733–758. https://doi.org/10.1016/B978-0-12-374546-0.00037-7

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F., Stevens, R., 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42, D206–D214. https://doi.org/10.1093/nar/gkt1226

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., Tyson, G.W., 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25, 1043–1055. https://doi.org/10.1101/gr.186072.114

Pinel, I.S.M., Kleikamp, H.B.C., Pabst, M., Vrouwenvelder, J.S., van Loosdrecht, M.C.M., Lin, Y., 2020. Sialic Acids: An Important Family of Carbohydrates Overlooked in Environmental Biofilms. Applied Sciences 10, 7694. https://doi.org/10.3390/app10217694

Pronk, M., Neu, T.R., van Loosdrecht, M.C.M., Lin, Y.M., 2017. The acid soluble extracellular polymeric substance of aerobic granular sludge dominated by

Defluviicoccus sp. Water Res 122, 148–158. https://doi.org/10.1016/j.watres.2017.05.068

Rivière, D., Desvignes, V., Pelletier, E., Chaussonnerie, S., Guermazi, S., Weissenbach, J., Li, T., Camacho, P., Sghir, A., 2009. Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. ISME J 3, 700–714. https://doi.org/10.1038/ismej.2009.2

Santander, J., Martin, T., Loh, A., Pohlenz, C., Gatlin, D.M., Curtiss, R., 2013. Mechanisms of intrinsic resistance to antimicrobial peptides of Edwardsiella ictaluri and its influence on fish gut inflammation and virulence. Microbiology (N Y) 159, 1471–1486. https://doi.org/10.1099/mic.0.066639-0

Santschi, P.H., Xu, C., Schwehr, K.A., Lin, P., Sun, L., Chin, W.-C., Kamalanathan, M., Bacosa, H.P., Quigg, A., 2020. Can the protein/carbohydrate (P/C) ratio of exopolymeric substances (EPS) be used as a proxy for their 'stickiness' and aggregation propensity? Mar Chem 218, 103734. https://doi.org/10.1016/j.marchem.2019.103734

Schmid, J., Sieber, V., Rehm, B., 2015. Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. Front Microbiol 6. https://doi.org/10.3389/fmicb.2015.00496

Schmidt, J.E., Ahring, B.K., 2000. Granular sludge formation in upflow anaerobic sludge blanket (UASB) reactors. Biotechnol Bioeng 49, 229–246. https://doi.org/10.1002/(SICI)1097-0290(19960205)49:3<229::AID-BIT1>3.0.CO;2-M

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Sekiguchi, Y., Takahashi, H., Kamagata, Y., Ohashi, A., Harada, H., 2001. In Situ Detection, Isolation, and Physiological Properties of a Thin Filamentous Microorganism Abundant in Methanogenic Granular Sludges: a Novel Isolate Affiliated with a Clone Cluster, the Green Non-Sulfur Bacteria, Subdivision I. Appl Environ Microbiol 67, 5740–5749. https://doi.org/10.1128/AEM.67.12.5740-5749.2001

Seviour, T., Derlon, N., Dueholm, M.S., Flemming, H.-C., Girbal-Neuhauser, E., Horn, H., Kjelleberg, S., van Loosdrecht, M.C.M., Lotti, T., Malpei, M.F., Nerenberg, R., Neu, T.R., Paul, E., Yu, H., Lin, Y., 2019. Extracellular polymeric substances of biofilms: Suffering from an identity crisis. Water Res 151, 1–7. https://doi.org/10.1016/j.watres.2018.11.020

Simate, G.S., Cluett, J., Iyuke, S.E., Musapatika, E.T., Ndlovu, S., Walubita, L.F., Alvarez, A.E., 2011. The treatment of brewery wastewater for reuse: State of the art. Desalination 273, 235–247. https://doi.org/10.1016/j.desal.2011.02.035

Song, M., Shin, S.G., Hwang, S., 2010. Methanogenic population dynamics assessed by real-time

quantitative PCR in sludge granule in upflow anaerobic sludge blanket treating swine wastewater. Bioresour Technol 101, S23–S28. https://doi.org/10.1016/j.biortech.2009.03.054

Sun, L., Toyonaga, M., Ohashi, A., Matsuura, N., Tourlousse, D.M., Meng, X.-Y., Tamaki, H., Hanada, S., Cruz, R., Yamaguchi, T., Sekiguchi, Y., 2016. Isolation and characterization of Flexilinea flocculi gen. nov., sp. nov., a filamentous, anaerobic bacterium belonging to the class Anaerolineae in the phylum Chloroflexi. Int J Syst Evol Microbiol 66, 988–996. https://doi.org/10.1099/ijsem.0.000822

Sun, X., Zhang, J., 2021. Bacterial exopolysaccharides: Chemical structures, gene clusters and genetic engineering. Int J Biol Macromol 173, 481–490. https://doi.org/10.1016/j.ijbiomac.2021.01.139

Talari, A.C.S., Martinez, M.A.G., Movasaghi, Z., Rehman, S., Rehman, I.U., 2017. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. Appl Spectrosc Rev 52, 456–506. https://doi.org/10.1080/05704928.2016.1230863

van Lier, J.B., Tilche, A., Ahring, B.K., Macarie, H., Moletta, R., Dohanyos, M., Hulshoff Pol, L.W., Lens, P., Verstraete, W., 2001. New perspectives in anaerobic digestion. Water Science and Technology 43, 1–18. https://doi.org/10.2166/wst.2001.0001

van Lier, J.B., van der Zee, F.P., Frijters, C.T.M.J., Ersahin, M.E., 2015a. Celebrating 40 years anaerobic sludge bed reactors for industrial wastewater treatment. Rev Environ Sci Biotechnol 14, 681–702. https://doi.org/10.1007/s11157-015-9375-5

van Lier, J.B., van der Zee, F.P., Frijters, C.T.M.J., Ersahin, M.E., 2015b. Development of Anaerobic High-Rate Reactors, Focusing on Sludge Bed Technology. https://doi.org/10.1007/10_2015_5012

Varki, A., Cummings, R.D., Aebi, M., Packer, N.H., Seeberger, P.H., Esko, J.D., Stanley, P., Hart, G., Darvill, A., Kinoshita, T., Prestegard, J.J., Schnaar, R.L., Freeze, H.H., Marth, J.D., Bertozzi, C.R., Etzler, M.E., Frank, M., Vliegenthart, J.F., Lütteke, T., Perez, S., Bolton, E., Rudd, P., Paulson, J., Kanehisa, M., Toukach, P., Aoki-Kinoshita, K.F., Dell, A., Narimatsu, H., York, W., Taniguchi, N., Kornfeld, S., 2015. Symbol Nomenclature for Graphical Representations of Glycans. Glycobiology 25, 1323–1324. https://doi.org/10.1093/glycob/cwv091

Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N.H., Prestegard, J.H., 2022. Essentials of Glycobiology [internet].

Veiga, M.C., Jain, M.K., Wu, W., Hollingsworth, R.I., Zeikus, J.G., 1997. Composition and role of extracellular polymers in methanogenic granules. Appl Environ Microbiol 63, 403–407.

Whitfield, C., Wear, S.S., Sande, C., 2020. Assembly of Bacterial Capsular Polysaccharides and

Exopolysaccharides. Annu Rev Microbiol 74, 521–543. https://doi.org/10.1146/annurev-micro-011420-075607

Wu, J., Jiang, B., Kong, Z., Yang, C., Li, L., Feng, B., Luo, Z., Xu, K.-Q., Kobayashi, T., Li, Y.-Y., 2021. Improved stability of up-flow anaerobic sludge blanket reactor treating starch wastewater by pre-acidification: Impact on microbial community and metabolic dynamics. Bioresour Technol 326, 124781. https://doi.org/10.1016/j.biortech.2021.124781

Xia, Y., Ju, F., Fang, H.H.P., Zhang, T., 2013. Mining of Novel Thermo-Stable Cellulolytic Genes from a Thermophilic Cellulose-Degrading Consortium by Metagenomics. PLoS One 8, e53779. https://doi.org/10.1371/journal.pone.0053779

Yamada, T., Sekiguchi, Y., 2018. Anaerolineaceae, in: Bergey's Manual of Systematics of Archaea and Bacteria. Wiley, pp. 1–5. https://doi.org/10.1002/9781118960608.fbm00301

Yamada, T., Sekiguchi, Y., Imachi, H., Kamagata, Y., Ohashi, A., Harada, H., 2005. Diversity, Localization, and Physiological Properties of Filamentous Microbes Belonging to Chloroflexi Subphylum I in Mesophilic and Thermophilic Methanogenic Sludge Granules. Appl Environ Microbiol 71, 7493–7503. https://doi.org/10.1128/AEM.71.11.7493-7503.2005

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., Xu, Y., 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 40, W445–W451. https://doi.org/10.1093/nar/gks479

Zhu, X., Kougias, P.G., Treu, L., Campanaro, S., Angelidaki, I., 2017. Microbial community changes in methanogenic granules during the transition from mesophilic to thermophilic conditions. Appl Microbiol Biotechnol 101, 1313–1322. https://doi.org/10.1007/s00253-016-8028-0

Supplemental Information

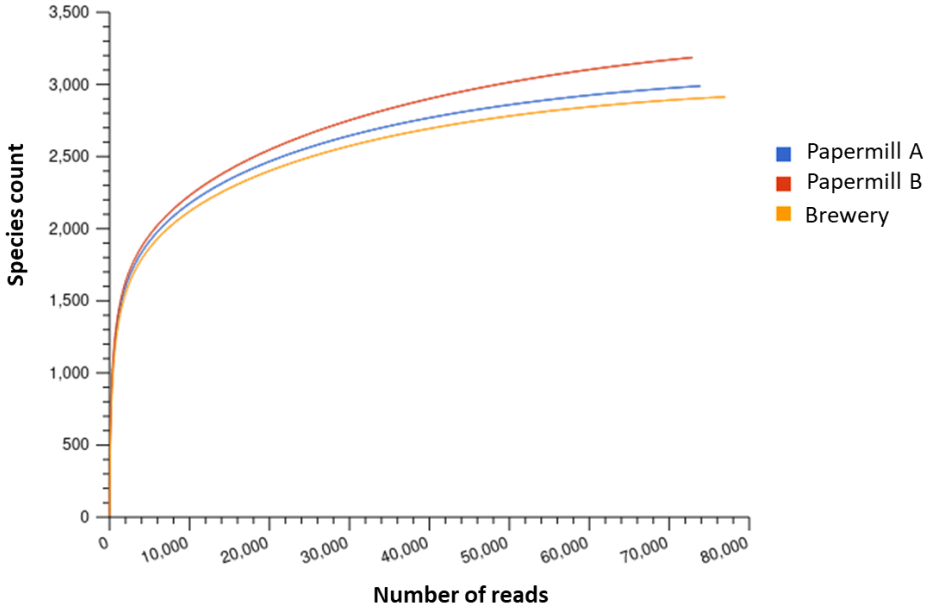


Figure S1. Rarefaction curve of the sequenced reads and species count in three metagenome samples: Papermill A, Papermill B, and Brewery.

Table S1. FT-IR absorbance bands assignments (Chirman and Pleshko, 2021; Lotti et al., 2019; Pronk et al., 2017; Talari et al., 2017).

Band number	Wavenumber (cm^{-1})	Functional assignment
1	3284-79	Carbohydrates, proteins, O-H symm. stretch, Amide A
2	3070-60	Proteins, aromatic amino acids
3	2960	Lipids, asymm. stretch CH_3
4	2924	Fatty acids, asymm. stretch CH_2
5	2850	Lipids, C-H stretching
6	1680-1620	Proteins, amide I
7	1560-20	Proteins, amide II
8	1456-2	Mostly lipids, proteins, CH_2 bending
9	1390	Fatty acids and proteins, CH_2 asymm. bending, COO^- stretching, respectively
10	1312	Proteins, amide III
11	1240	Phosphate, Sulfate and proteins, PO_2 asymm. Stretch, sulfate half ester, Amide III C-N stretch
12	1100-932	Polysaccharides, oligosaccharides and monosaccharides, asymm. stretch C-O-C, C-OH stretch

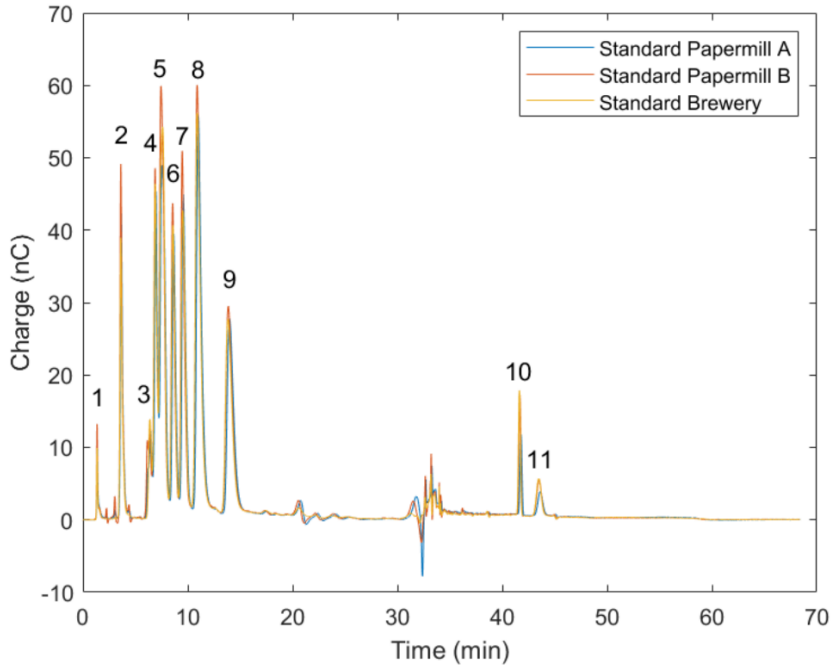


Figure S2**.** HPAE-PAD chromatograms of the sugar standards used.

Table S2. Measured retention times of sugar standards in the HPAE-PAD analysis with corresponding peaks in sample chromatogram.

Peak number	**Sugar standard**	**Ret. Time (min)**	**Corresponding peak**
1	glycerol	1.3	1
2	fucose	3.6	5
3	rhamnose	6.3	7
4	glucosamine	6.9	8
5	arabinose	7.4	9
6	galactose	8.6	10
7	glucose	9.	11
8	mannose	10.87	12
9	ribose	13.94	13
10	galacturonic	41.64	n.a.
11	glucuronic	43.64	n.a.

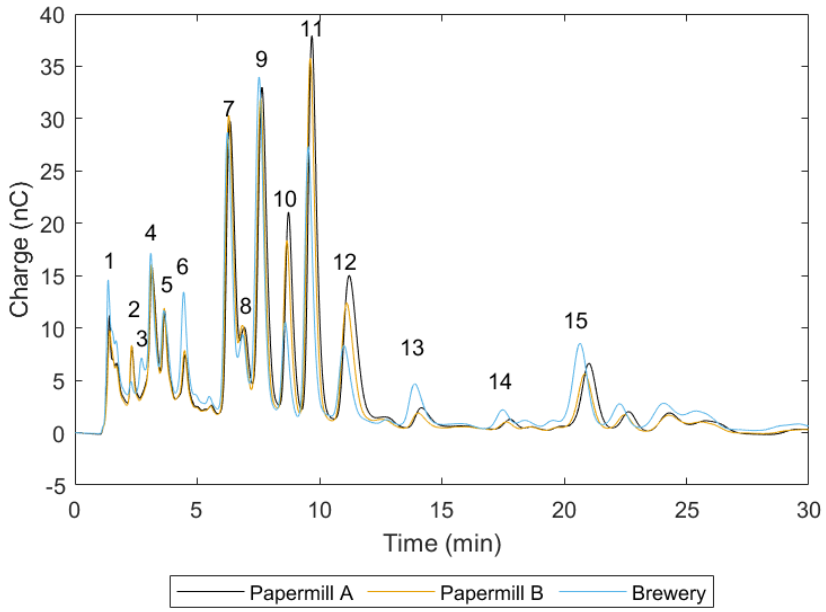


Figure S3. HPAE-PAD chromatograms for the EPS of the granular sludge samples.

Table S3. Raw data for the HPAE-PAD data portrayed in the Figure 2.

Sugar monomer	**Papermill A (% of total EPS)**	**Papermill B (% of total EPS)**	**Brewery (% of total EPS)**
Fucose	0,09%	0,08%	0,00%
Galactosamine	0,77%	0,71%	0,96%
Rhamnose	0,40%	0,55%	0,44%
Glucosamine	0,40%	0,93%	0,58%
Galactose	0,25%	0,33%	0,07%
Glucose	1,91%	2,09%	1,43%
Mannose	1,02%	1,11%	0,54%
Galacturonic acid	3,24%	4,82%	4,07%
Glucuronic acid	0,92%	1,54%	2,03%
Total	10,85%	12,17%	10,12%

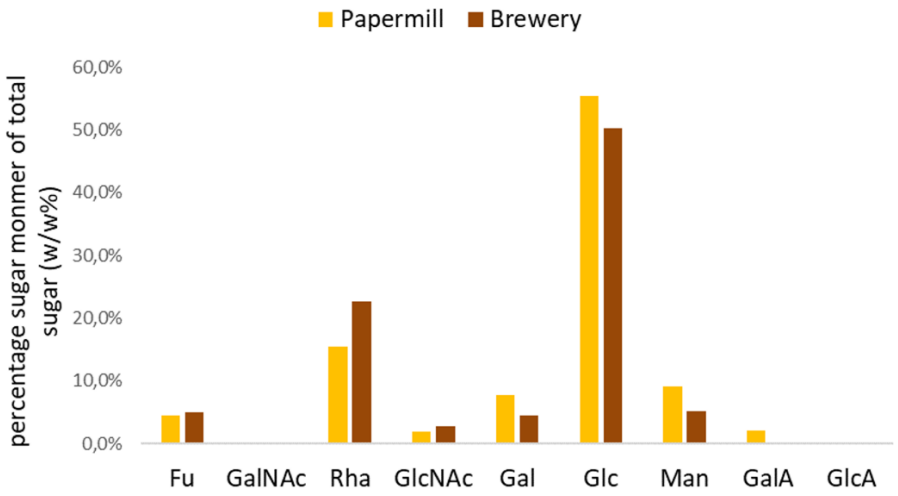


Figure S4. Glycosyl composition analysis with TMS-derivatization GC-MS analysis. Abbreviations are according to symbol nomenclature for glycans (Varki et al., 2015). Fu stands for fucose, GalNAc for N-acetyl galactosamine, Rha for rhamnose, GlcNAc for N-acetyl glucosamine, Gal for galactose, Glc for glucose, Man for mannose, GalA for galacturonic acid and GlcA for glucuronic acid.

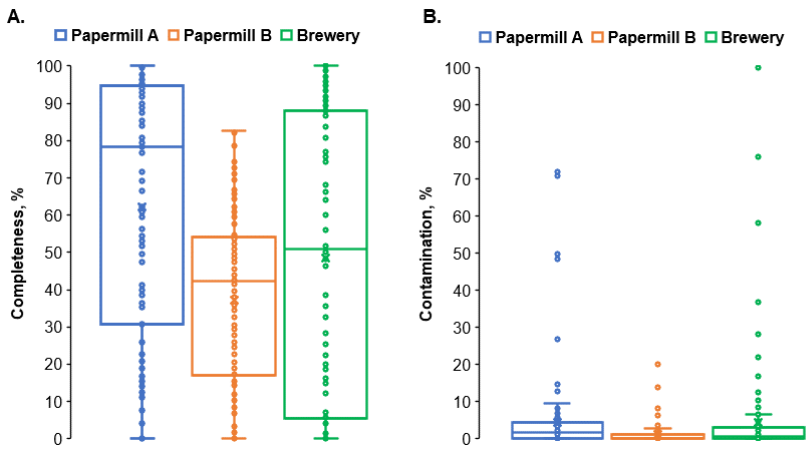


Figure S5. Completeness (**A**) and contamination (**B**) of the metagenome assembled bins from the three sludge samples: Papermill A, Papermill B, Brewery.

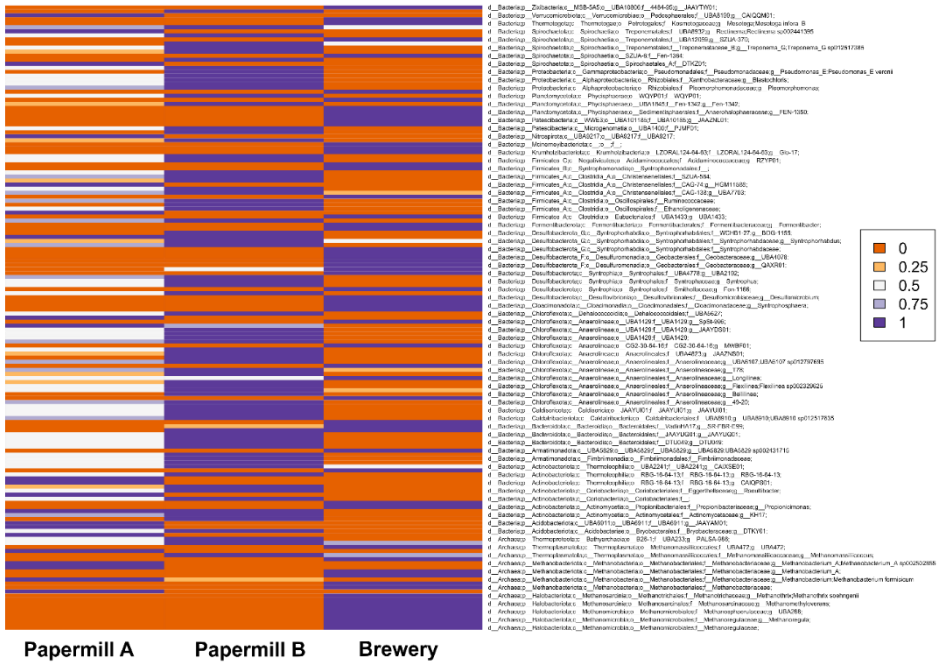


Figure S6. Normalized abundance and phylogenetic placement of all the MAGs from the three granular sludge samples.

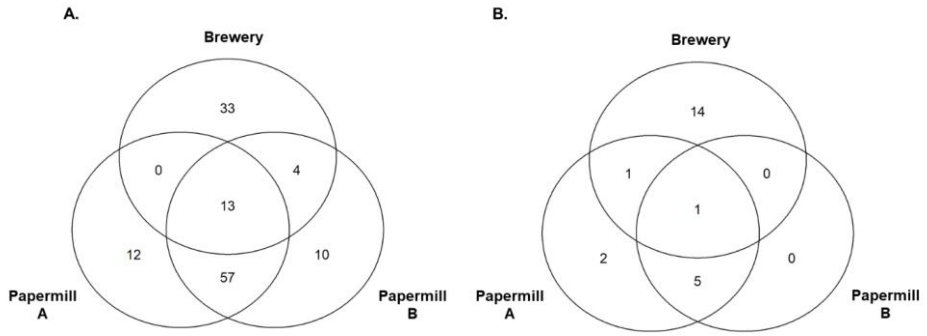


Figure S7. Distribution of shared and unique taxa from quality filtered MAGs of the three sludge samples. (**A**.) All Bacterial MAGs with defined taxonomic classification; (**B.**) All Archaeal MAGs with defined taxonomic classification.

Table S4. The molecular functions of the annotated genes within the three anaerobic sludge sample based on the KEGG Orthology. Numbers in each cell represent the number of mapped genes, while the color gradient goes from green (high number of genes) to red (low number of genes).

	Papermill A	Papermill B	Brewery
Amino acid metabolism	16613	17319	16065
Carbohydrate metabolism	12713	13262	12156
Lipid metabolism	1995	2109	1793
Nucleotide metabolism	4338	4405	4196
Energy metabolism	5227	5421	5993
Biosynthesis of other secondary metabolites	1441	1516	1272
Glycan biosynthesis and metabolism	1512	1642	1653
Metabolism of cofactors and vitamins	5450	5922	6018
Metabolism of terpenoids and polyketides	954	993	908
Xenobiotics biodegradation and metabolism	500	594	343
Cell communication	10	12	21
Cell growth and death	1740	1876	1662
Cell motility	2530	3027	1743
Transport and catabolism	746	766	763
Folding, sorting and degradation	2536	2773	2766
Replication and repair	4301	4736	4177
Transcription	1366	1463	1324
Translation	8983	9533	9006
Membrane transport	9995	10775	8988
Signal transduction	4522	4945	3871
Signaling molecules and interaction	6	5	6
Environmental adaptation	216	256	192
Excretory system	2	1	9

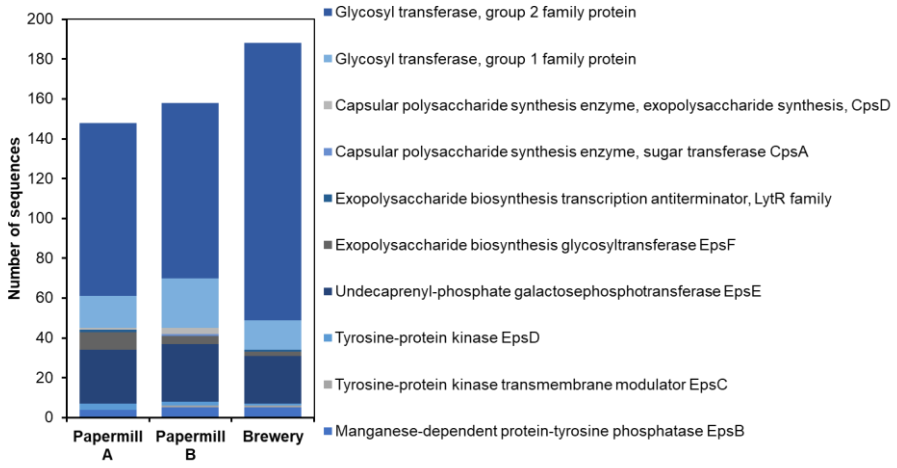


Figure S8. Distribution of genes identified in the three sludge samples related to the exopolysaccharide synthesis.

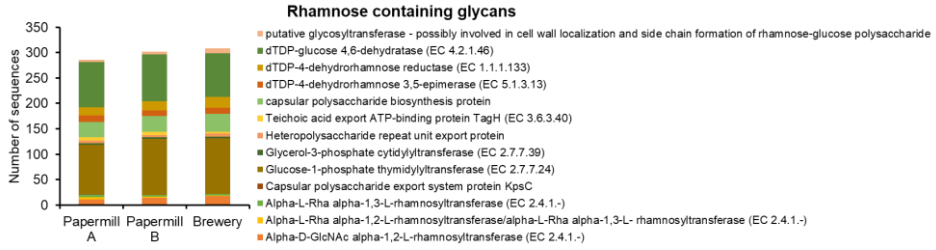


Figure S9. Distribution of genes identified in the three sludge samples related to the metabolism of rhamnose containing glycans.

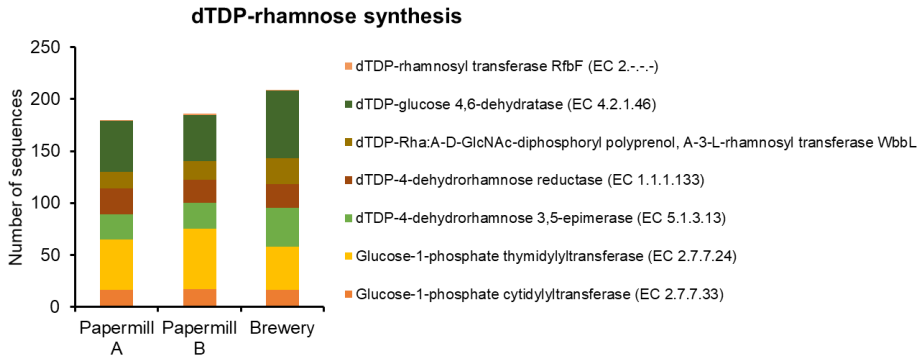


Figure S10. Distribution of genes identified in the three sludge samples related to the dTDP-rhamnose synthesis.

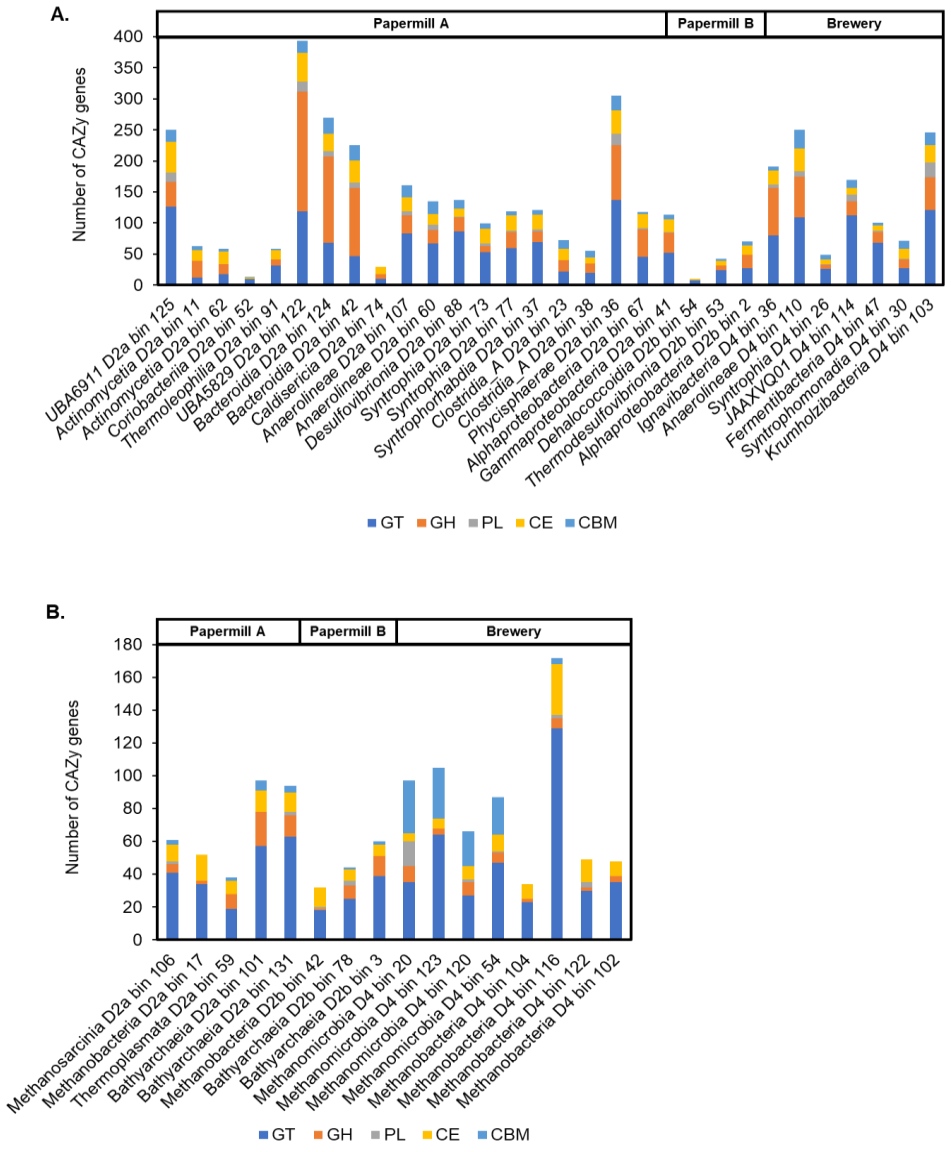


Figure S11. Characterization of the CAZy-coding genes in the qualified MAGs (more than 70% completeness, less than 5% contamination) from the three sludge samples. MAGs classified as Bacteria (A) and as Archaea (B) are both shown. GT – Glycosyl Transferases, GH – Glycoside Hydrolases, PL – Polysaccharide Lyases, CE – Carbohydrate Esterases, CBM – Carbohydrate-Binding Modules.

References

Chirman, D., Pleshko, N., 2021. Characterization of bacterial biofilm infections with Fourier transform infrared spectroscopy: a review. Appl Spectrosc Rev 56, 673–701. https://doi.org/10.1080/05704928.2020.1864392

Lotti, T., Carretti, E., Berti, D., Martina, M.R., Lubello, C., Malpei, F., 2019. Extraction, recovery and characterization of structural extracellular polymeric substances from anammox granular sludge. J Environ Manage 236, 649–656. https://doi.org/10.1016/j.jenvman.2019.01.054

Pronk, M., Neu, T.R., van Loosdrecht, M.C.M., Lin, Y.M., 2017. The acid soluble extracellular polymeric substance of aerobic granular sludge dominated by Defluviicoccus sp. Water Res 122, 148–158. https://doi.org/10.1016/j.watres.2017.05.068

Talari, A.C.S., Martinez, M.A.G., Movasaghi, Z., Rehman, S., Rehman, I.U., 2017. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. Appl Spectrosc Rev 52, 456–506. https://doi.org/10.1080/05704928.2016.1230863

Varki, A., Cummings, R.D., Aebi, M., Packer, N.H., Seeberger, P.H., Esko, J.D., Stanley, P., Hart, G., Darvill, A., Kinoshita, T., Prestegard, J.J., Schnaar, R.L., Freeze, H.H., Marth, J.D., Bertozzi, C.R., Etzler, M.E., Frank, M., Vliegenthart, J.F., Lütteke, T., Perez, S., Bolton, E., Rudd, P., Paulson, J., Kanehisa, M., Toukach, P., Aoki-Kinoshita, K.F., Dell, A., Narimatsu, H., York, W., Taniguchi, N., Kornfeld, S., 2015. Symbol Nomenclature for Graphical Representations of Glycans. Glycobiology 25, 1323–1324. https://doi.org/10.1093/glycob/cwv091

Chapter 5: FT-IR micro-spectroscopy for imaging the extracellular matrix composition in biofilms

Abstract

Microorganisms form granules by embedding themselves in an extracellular matrix through the secretion of extracellular polymeric substances (EPS). The extracellular matrix is a complex structure comprising of e.g. proteins, polysaccharides, lipids, and extracellular DNA. Understanding the function of individual EPS components within the matrix requires not only information on the composition of the extracellular matrix, but also information on the spatial distribution. Advanced imaging techniques have been proposed and implemented for the visualization of the extracellular matrix, but are often to detailed. Untargeted approaches like FT-IR micro-spectroscopy would allow for a broader exploration of the polymer distribution. In this study FT-IR micro-spectroscopy analysis was implemented on sliced anaerobic granular sludge to explore the EPS distribution. Visualization of absorbance heights of single wavenumbers showed there was a higher polysaccharide content in the EPS at the outer edge of the granule, shifting to a higher protein concentration toward the centre, with a boundary approximately 150 μm from the surface, which was in accordance with layer of fermentative bacteria described in literature. The complexity of polymers inside the extracellular matrix meant that many functional groups were overlapping, making FT-IR annotation challenging. To address this, principal component analysis and two-dimensional correlation spectroscopy analysis were included in the analysis. These methods enabled the identification of overlapping functional groups and correlations between functional groups. Positive correlations between protein and polysaccharide functional groups suggested the presence of glycoproteins, which has been regularly described in chemical EPS analysis studies. Additionally, correlations between sulfated compounds and protein/polysaccharide functional groups indicated potential co-localization in the extracellular matrix. Differences in positive correlations of sialic acids with polysaccharides suggest variations in polysaccharide or other polymer compositions, possibly caused by differences in the microbial community.

Keywords

EPS, granular sludge, chemometrics, 2D-COS

Highlights

- Layered structure of microorganisms was reflected in the EPS distribution visualized with FT-IR micro-spectroscopy

- The combination of FT-IR micro-spectroscopy and chemometric analysis allowed for comprehensive insights into the polymer distribution

published as:

FT-IR micro-spectroscopy for imaging the extracellular matrix composition in biofilms. Stefan de Bruin, Carina Hof, Mark van Loosdrecht, Diana Z. Sousa, Yuemei Lin. bioRxiv 2024.08.26.609418; doi: https://doi.org/10.1101/2024.08.26.609418

1. Introduction

Using granular sludge in wastewater treatment makes it possible to remove organic pollutants at higher loading rates compared to flocculent sludge (Van Lier et al., 2015). Microorganisms form granules by secreting polymers that embed them in an extracellular matrix (Flemming & Wingender, 2010). Extracellular polymeric substances (EPS) are a mixture of polymers found in the extracellular matrix, and in general contains: proteins, polysaccharides, lipids, extracellular DNA and combinations of these polymers (Flemming & Wingender, 2010; Seviour et al., 2019). EPS provide the structural support, and thus influence matrix stability (Felz et al., 2016; Bou-Sarkis et al., 2022). This structural support is caused by stabilization through either covalent bonding between polymers, hydrogen bonding or through cross-linking of charged groups in the EPS (Costa et al., 2018; Lotti et al., 2019). Microorganisms adapt to environmental cues by secreting charged moieties that influence the diffusion of ions into the extracellular matrix (de Graaff et al., 2019; Gagliano et al., 2018; Li et al., 2018). Amongst the most acidic charged moieties are sialic acids, phosphorylated and sulfated compounds. These charged groups were found to be present in the EPS of various biofilm systems (Boleij et al., 2020; de Bruin et al., 2022; Chen et al., 2023; Felz et al., 2020; Bourven et al., 2015). In both aerobic and anaerobic granules it was seen that sialic acids were more prevalent in high molecular weight fractions, whereas sulfated compounds were present throughout the whole molecular weight range (Chen et al., 2023). Whether this influences the functionality is not yet known, as the function of a specific polymer inside the extracellular matrix is also dependent on the location.

Seviour et al., (2019) proposed a roadmap for understanding the role of individual EPS components inside the extracellular matrix. Two routes have been suggested for imaging biofilms. The first is to extract structural EPS and determine the specific polymer structure, through chemical analysis and use this information to find specific probes for these chemical structures (Figure 1). Traditionally this is done by extracting the EPS from the biofilm using e.g. alkaline EPS extraction, and then analysing it using colorimetric and chemical analysis (Felz et al., 2019). Examples of these chemical analysis are nuclear magnetic resonance monomer analysis and high performance anion exchange chromatography with pulsed amperometric detection to identify polysaccharides and liquid chromatography combined with tandem mass spectrometry for protein identification (Gonzalez-Gil et al., 2015; Dubé and Guiot, 2019). The identified polymers can then be imaged by targeted imaging techniques like e.g. fluorescent lectin binding analysis for glycans and antibody for proteins (Neu & Kuhlicke, 2022). The advantage of this route is that specific compounds can be localized with a small optical resolution, making it possible to image the EPS of single cells inside the biofilm. The second route is by using untargeted imaging techniques like e.g. MALDI-TOF MS, Raman micro-spectroscopy and FT-IR micro-spectroscopy (Gowen et al., 2015). With these techniques, the spatially resolved spectra offer more chemical information than can be generated with typical microscopy. As such, these techniques detect a broad range of polymers that are measured simultaneously. The function of a specific polymer relies on the presence

of co-localized polymers/monomers. Thus, in comparison to the targeted imaging route, the untargeted imaging route provides broader imaging capabilities.

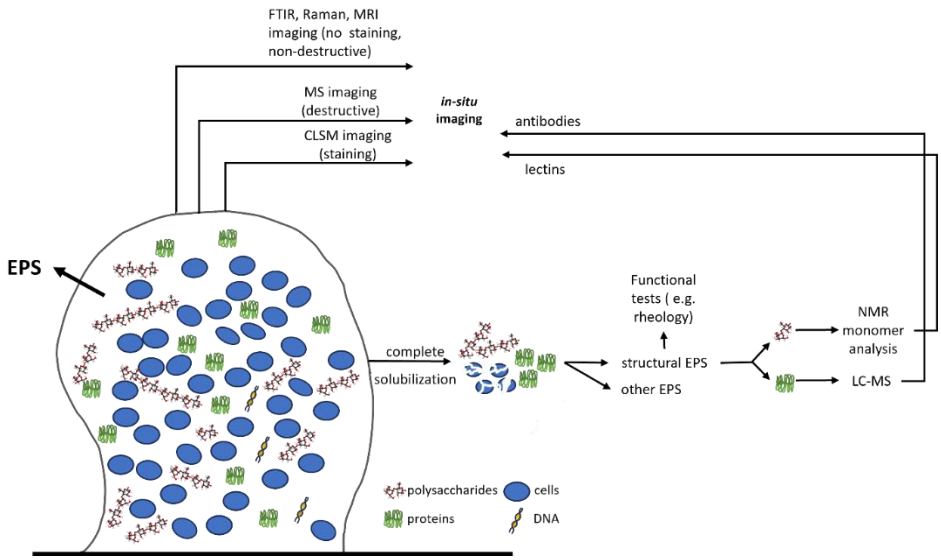


Figure 1. Roadmap adapted from Seviour et al. (2019), showcasing potential routes for imaging the biofilm

With FT-IR technique, many different functional groups present in polymers can be detected because vibrational spectra are molecule specific. FT-IR imaging is a user-friendly and non-destructive technique with minimal sample preparation and a high signal to noise ratio (Gowen et al., 2015). Mapping of the polymer distributions can be achieved by using the absorbance height, or area, of bands associated with specific polymer types (Jiang & Rieppo, 2006). However, absorbance bands of functional groups often correspond to multiple polymer types, leading to spectral complexity. In these cases, it is beneficial to utilize all the information stored in the spectrum simultaneously. This task is commonly addressed with chemometrics, a field that applies mathematical and statistical methods to analyse chemical data. A simple way of reducing the complexity of the spectral data is by reducing the dimensionality. With principal component analysis (PCA) this data dimensionality reduction is achieved while retaining the majority of the information. PCA describes the absorbance spectra as a linear combination of component spectra (principal components) and their scores. By imaging the principal component scores, it is possible to visualize the polymer distribution with limited loss of information (Gowen et al., 2015; Chirman & Pleshko, 2021). To see if polymers co-localize in the extracellular matrix the corelation of different functional groups can be explored. Generalized two-dimensional correlation spectroscopy (2D-COS) is a way to analyse correlations in the spectra between absorbance bands under the influence of e.g. time, pH, temperature or space (Jiang & Rieppo, 2006; Noda et al., 2000). This technique is used to determine the correlation spectrum of spatially resolved

absorbance spectra (Lasch & Noda, 2019). If the functional groups are correlated, it indicates those polymers are co-localized.

The aim of this study is to explore the extracellular matrix polymer composition using FT-IR micro-spectroscopy. Through the use of chemometrics a more guided effort can be achieved of the untargeted approach to visualize the polymer composition in the extracellular matrix. The use of FT-IR micro-spectroscopy may enhance our understanding of the spatial arrangement of polymers in the granular sludge matrix, providing valuable insights into its composition and structure.

2. Materials and methods

2.1. Sample handling and FT-IR screening of EPS absorbance bands

Granular sludge was collected from a full scale anaerobic wastewater treatment plant treating papermill effluent similar to other studies before (Chen et al., 2023; Doloman et al., 2024). The first step in the analysis of the extracellular matrix is to identify the functional groups present in the FT-IR absorbance spectrum of EPS. EPS extraction was performed on washed and lyophilized granular sludge using an alkaline extraction method as described earlier (Pinel et al., 2020). FT-IR spectra of the extracted EPS were recorded with a FT-IR spectrophotometer (Spectrum 100, PerkinElmer, Shelton, USA) over a wavenumber range of 4000 cm^{-1} to 600 cm^{-1} with 16 accumulations and 4 cm^{-1} spectral resolution.

2.2. Homogeneous granule slicing and data acquisition

FT-IR micro-spectroscopy using an ATR-imaging crystal is a technique that measures the absorbance spectra at the surface of the ATR crystal. Therefore, if unsliced granular sludge is used for the imaging analysis, the external surface can be imaged. Slicing of the granular sludge ensures the possibility of imaging the internal granular structure. Before slicing, granules of 1 mm in diameter were selected and were embedded in an optimal cutting temperature medium (Neg-50™ Frozen section media, Epredia™). Granules were sliced with a cryotome to obtain slices that were 15 μm in depth (cabinet temperature -20°C, mount temperature -25°C; Leica CM1900).

FT-IR micro-spectroscopy imaging was performed in ATR imaging mode with a germanium attenuated total reflectance crystal (Spotlight 400 FTIR microscope, PerkinElmer, Shelton, USA). Absorbances were recorded in the 4000 cm^{-1} to 600 cm^{-1} range with 4 accumulations, a spatial resolution of 1.56 μm^2 and a spectral resolution of 4 cm^{-1}. The measured area could be set to be anywhere between 50 x 50 μm and 500 x 500 μm, and was set to 300 x 300 μm for this study.

2.3. Data visualization

The workflow for FT-IR imaging as employed in this study is shown in figure 2. The steps are discussed in detail in the following text, for additional information how this is applied see the supplemental MATLAB script (Supplemental information S1).

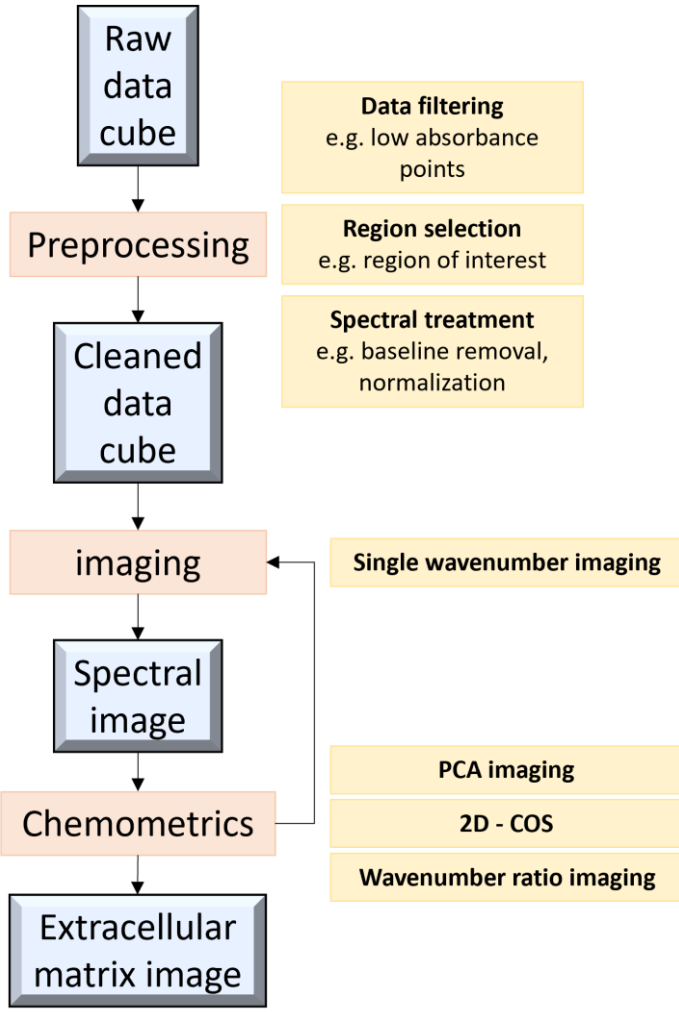


Figure 2. workflow diagram FT-IR imaging data pipeline

2.3.1. Data pre-processing

For the acquisition of the data it was important to ensure good sample contact between the sample and imaging crystal. This would make it possible to obtain data with a high signal over noise ratio. After acquisition the data was smoothed with a Savitsky-golay filter, and base-line was corrected by a zero-order baseline correction. Subsequently normalization was performed by feature scaling. Due to the heterogeneity of the biofilm, there are discrepancies on the average absorbance height between locations on the granule slice. This can partly be solved by ensuring slicing as homogeneously as possible. Supplemental Figures S1-S2 show the microscopic images of the sliced granular sludge. The grooves of the metal disc on which the samples were mounted affect the absorbance height (Supplemental Figure S3). The average absorbance height images were used to locate and filter areas with

a high signal to noise ratio. Based on the data, pixels with an average absorbance less than 0.3 were discarded (Supplemental Figures S4-S5).

2.3.2. Single absorbance band imaging

The single point absorbance imaging was created by mapping the absorbance height of proteins (Amide I band at 1630 cm^{-1}), polysaccharides (C-O-C band at 1030 cm^{-1}), lipids (C–H stretch of > CH_2 groups, dominated by lipid contributions at 2920 cm^{-1}) (Talari et al., 2017), sialic acids (C = O stretch of ester specifically correlated to sialic acids at 1726 cm^{-1}) (Nallala et al., 2020) and putative sulfate groups (sulfate stretch at 1206 cm^{-1}) (Brézillon et al., 2014). The absorbance height for all the given wavenumbers was then scaled and mapped across the measured area to show the relative differences in distribution.

2.3.3. Principal component analysis

The interpretation of large datasets, like those from FT-IR imaging measurements, is challenging. Interpretation of data is increasingly complex in FT-IR absorbance spectra due to the presence of broad and convoluted absorbance bands. To address this, PCA is employed, which reduces highly dimensional data to a more manageable format, while conserving as much of the variability of the data as possible. The application of PCA results in the deconstruction of the absorbance signal into two matrices: the loadings matrix and the scoring matrix. The loadings matrix assigns weights to each wavenumber for every principal component, while the scoring matrix assigns scores to each absorbance spectrum for every principal component. The first principal component captures the highest amount of variability. Subsequent principal components are independent of each other, each describing the most variability in turn. In this study, the PCA loadings and scores were calculated using the built-in function in MATLAB. This methodology enables a more accessible and insightful understanding of the intricate FT-IR data, facilitating the extraction of meaningful patterns and information from the complex spectra (Jolliffe & Cadima, 2016).

2.3.4. Two-dimensional correlation spectroscopy

Generalized two-dimensional correlation spectroscopy is a powerful analytical method for the interpretation of vibrational spectra that is known primarily from use in NMR spectroscopy. Its use however, is not limited to NMR and can be used in other vibrational spectroscopy techniques as well (Noda et al., 2000). Variations in the spectrum arising from, in this case, spatial differences can serve as input for the correlational analysis. The synchronous 2D correlation spectrum essentially mirrors the covariance matrix. The matrix is derived by calculating the covariance between all combinations of spectral feature vectors at the specified spectral variables v_1 and v_2. By applying color-coding to represent the covariance values at v_1 and v_2 a 2D covariance map can be generated (Lasch & Noda, 2019).This map visually encapsulates the interrelationships among spectral features and provides an overview of covariations across FT-IR spectra.

3. Results

3.1. Determination of the FT-IR absorbance regions of interest

An initial screening of the functional groups that can be found in the EPS was performed using FT-IR on extracted EPS. There is overlap between the absorbance bands of functional groups. Therefore analysing anaerobic granular sludge or any other heterogeneous biofilm is challenging. To get an understanding of what absorbance bands are important in EPS, first an extracted EPS sample was analysed. The EPS appears to be protein dominant as shown in the height of the protein region absorbance bands. The relatively high and wide bands in the polysaccharide region indicate a high polysaccharide content that are likely present in large polymer chains. Other functional groups include absorbance bands annotated to the lipid region and anionic compounds like phosphate, sulphate and sialic acids (Figure 3).

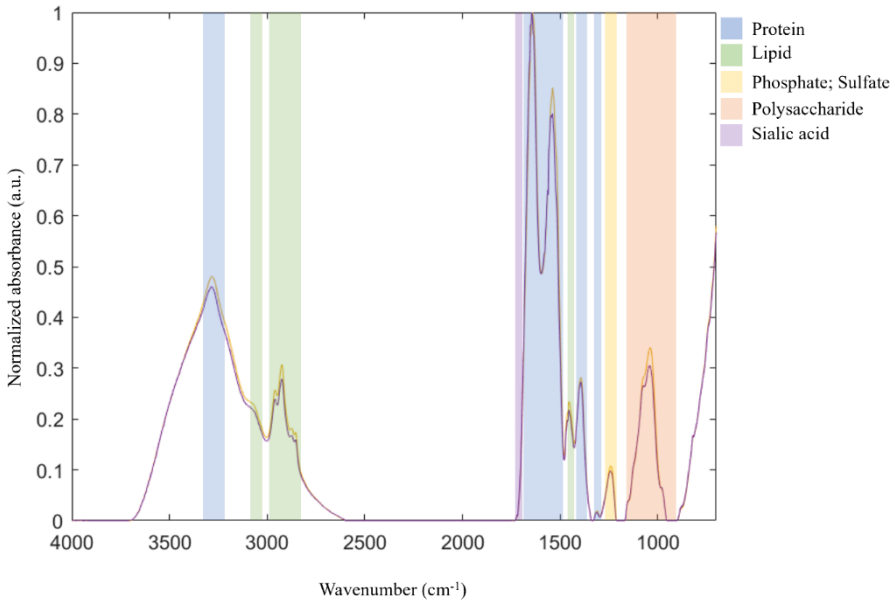


Figure 3. FT-IR absorbance spectra of EPS extracted from anaerobic granular sludge with annotated polymer functional region absorbance bands

3.2. Single-point absorbance

The chemical distribution of the most prevalent components in EPS were mapped (Figure 4) for a sliced granule. Firstly, it can be seen that, all the polymers and sialic acids are unevenly distributed in granules with different intensity (the intensity is increasing from green to red), indicating the EPS composition of the granule is heterogeneous. e.g. in the protein chemical image, it can be seen that the bottom left corner is more protein rich relative to other mapped areas (indicated by arrow). Similarly, there are specific areas where polysaccharides, lipids, sialic acids and sulfated/phosphorylated polymers are more abundant than at other areas (indicated by the arrows at the specific chemical image). Secondly, there are both positive and

negative correlations between the signals of different components i.e. polysaccharides and proteins signals are in general positively correlated except for the top and top right zone. The lipid rich regions were distributed similarly to the protein rich regions. Sulfated/Phosphorylated polymers followed similar abundancy profiles as polysaccharides and lipids (as indicated in 4B and 4C) except for the area indicated by the arrow in 4E. In comparison, the sialic acids signal seemed to be negatively correlated with the signal of other components.

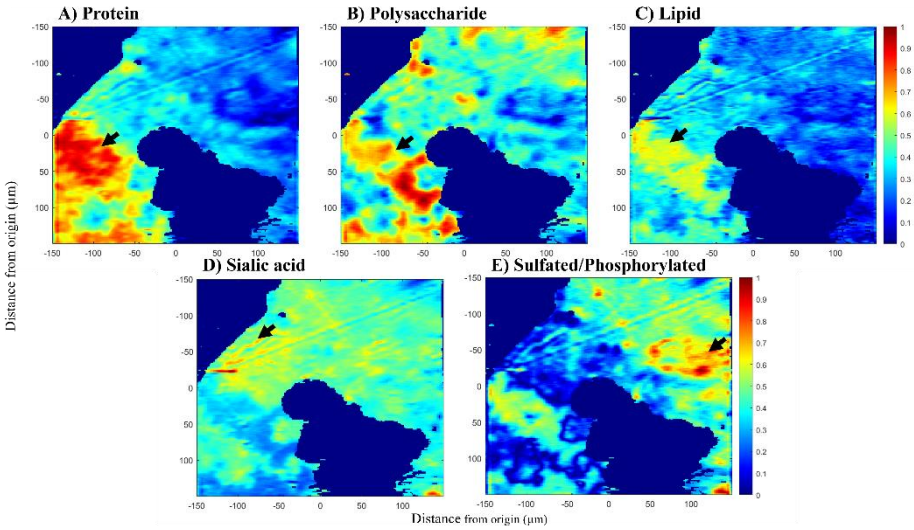


Figure 4. Scaled mapping of FT-IR functional groups absorbance band heights of sliced anaerobic granular sludge treating papermill wastewater. Protein) Height of the protein region absorbance band at 1630 cm^{-1}. Polysaccharide) Height of the polysaccharide region absorbance band at 1030 cm^{-1}. Lipid) Height of the lipid region absorbance band at 2920 cm^{-1}. Sialic acid) Height of the sialic acid region absorbance band at 1726 cm^{-1}. Sulfate) Height of the putative sulfate region absorbance band at 1206 cm^{-1}. The origin in this picture is the centre of the ATR imaging crystal, and so also the centre of the imaged area. Regions that were filtered out due to weak signal are shown in dark blue. The black arrows indicate regions of high absorbance.

3.3. Principal component analysis for dimension reduction

Proteins, polysaccharides, lipids and sialic acids contain multiple functional groups. The signal of these functional groups can overlap with each other which makes it difficult to get a complete characterization by only looking at single wavenumber absorbance bands for each component. In order to visualize the biggest differences in the data set with the complex combinations of functional groups whilst still maintaining most of the information at a reasonable level, principal component analysis (PCA) was applied. Figure 5 shows the spectrum of the second, third and fourth principal component (PC) loadings calculated from the spectral data. While the first PC could explain most of the variability in the area (67.9%), it showed the biggest variety in the data could be explained by the absorbance intensity of most functional groups. To image how more specific compounds were distributed across

the granule slice, the second, third and fourth PC were chosen. These explained 23.6%, 3.6% and 1.3% of the variability in the data, respectively. The loading of PC2 was protein dominant and emphasized amide functional groups: 3300, 3070, 1660, 1560, 1500 and 1240 cm^{-1}. However, other functional groups like lipid CH_3: 2960 and 2900 cm^{-1}, polysaccharides: 1110, 1070 and 1000 cm^{-1} and sialic acid at 1728 cm^{-1}, where also present in the PC2 loading. The PC3 loading was more polysaccharide dominant and emphasized polysaccharide functional groups:1300, 1132 and 968 cm^{-1}. Other important functional groups were lipid CH_2 stretch: 2994 and 2832, sulfate/phosphate:1218 cm^{-1} and amide functional groups: 1676 and 1480 cm^{-1}. Lastly, PC4 was polysaccharide dominant with highest weight on the polysaccharide functional groups: 1100 and 1014 cm^{-1}. Additional functional groups were amide: 1684, 1668 and 1580 cm^{-1}, and CH_2 stretch 2920 cm^{-1}.

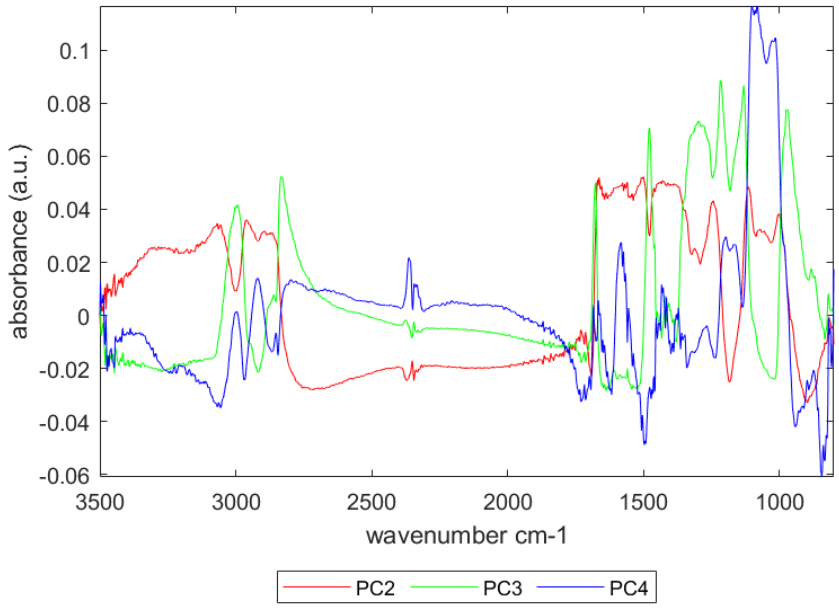


Figure 5. Principal components loading spectrum of a granular sludge slice.

The PC scores of the spatially resolved FT-IR spectra were mapped and are shown in figure 6. The left bottom part had a relatively high score for PC2, which coincided with the mappings of the protein and part of the polysaccharide, lipid and sulfate functional groups seen in figure 4. The region which the highest overlap of functional groups was marked in yellow indicating an overlap of PC2 and PC3 scores. Scores on PC3 were also high in the right region of the mapped area. This region had relatively high absorbance for sulfated/phosphorylated functional groups. High scores on PC4 indicated a more dominant presence of polysaccharides and the regions with high scores on PC4 also had relatively higher polysaccharide functional group absorbance in these areas. Mapping of the PC scores showed that

there are many overlapping functional groups, but that it was possible to distinguish between similar functional groups by looking functional group pairs found in the PCs.

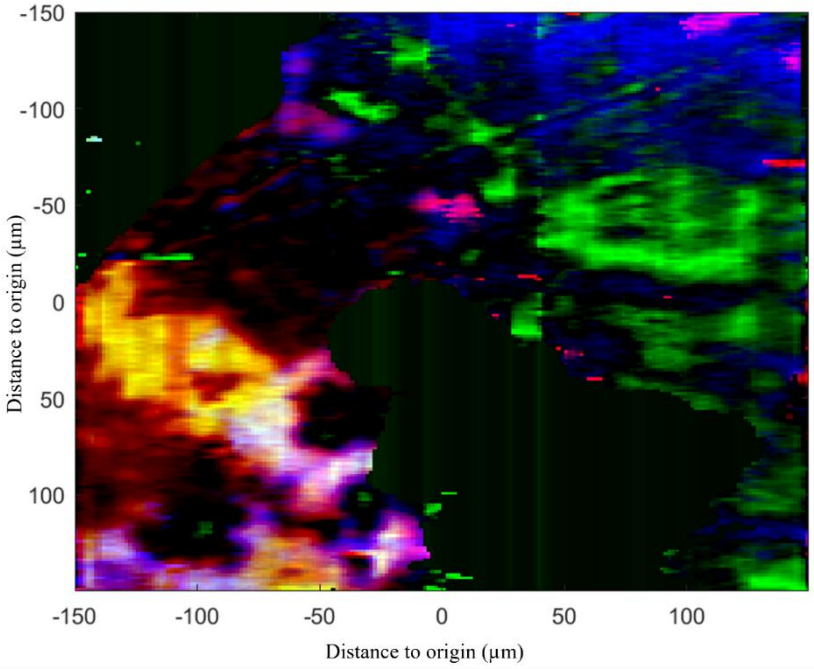


Figure 6. Mapping of the PC scores for PC2 (red), PC3 (green) and PC4 (blue) across the granule slice.

3.4. 2D-COS analysis

The EPS composition in anaerobic granules was spatially heterogeneous and there was strong overlap of bands from different components. In order to decipher the correlations between overlapping functional groups and detect compositional changes which are 'hidden' from PCA and single wavenumber absorbance, two-dimensional correlation spectroscopy (2D-COS) analysis was used. The spatially resolved FT-IR absorbance spectra were used to generate the synchronous correlation spectrum shown in figure 7. It is observed that, for each individual component, there are positive correlations among their specific typical absorbance regions along the scanned surface area. i.e. –For polysaccharides, absorbance at the regions of 850 – 920 cm^{-1}, 1000 – 1050 cm^{-1} and 1160 – 1200 cm^{-1} are positively correlated. While for proteins, amide regions at 1380 – 1460 cm^{-1} and 1500 – 1700 cm^{-1} were also positively correlated. The functional group annotated to sialic acid (1726 cm^{-1}) was also positively correlated along the scanned area.. For lipids, the only functional group that was positively correlated was for the CH_3 stretch found at

2840 – 2880 cm^{-1}. Secondly, in between different components, positive correlations were observed among lipid, polysaccharides and proteins at 2960 and 2860 cm^{-1} (-CH stretch, with lipid dominantly contributed), 1000 – 1050 cm^{-1} (C-O-C stretch) and 1380 – 1460 cm^{-1} and 1500 – 1700 cm^{-1}(amide bands). Interestingly, lipid seemed to be positively correlated with the sialic acid functional group (at 1726 cm^{-1}) only at 2920 cm^{-1} and not at other regions annotated to lipids. The sialic acid functional group was mainly correlated with the polysaccharide functional groups at 850 – 920 cm^{-1}, 1000 – 1050 cm^{-1} and 1160 – 1200 cm^{-1} regions. Proteins and polysaccharides were positively correlated at 1630 cm^{-1} and the 1000 – 1050 cm^{-1} spectral regions. Phosphorylated or sulfated polymers at 1234 cm^{-1} seemed to be mainly positively correlated with the polysaccharide and protein functional groups at 1000 – 1050 cm^{-1} and 1500 – 1700 cm^{-1}, respectively. Another spectral band appeared at 1180 cm^{-1} was observed positively correlated with the sialic acid band at 1726 cm^{-1}, but not with any other absorbance.

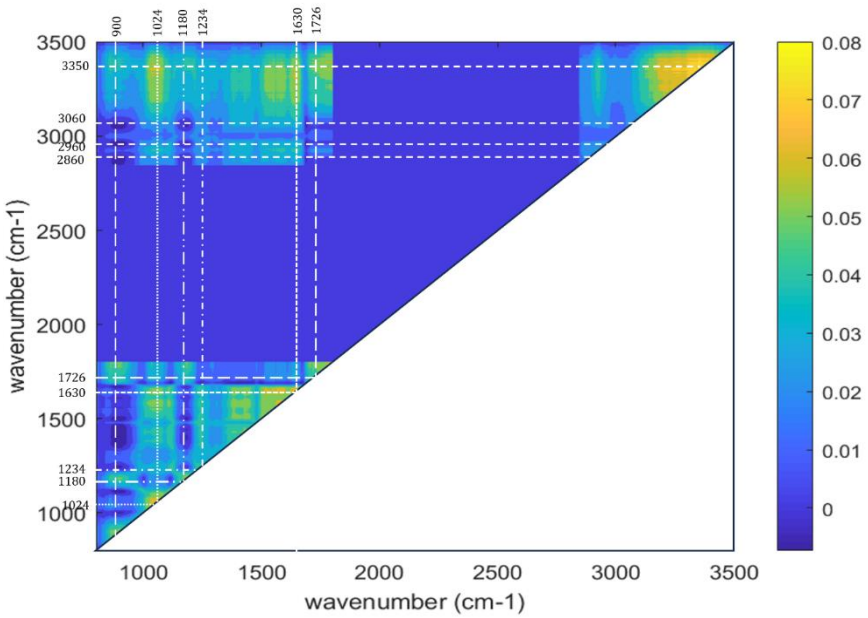


Figure 7. Generalized synchronous two-dimensional correlation spectrum contour plot derived from the spatially resolved spectra.

3.5. Imaging at the outer edge of the anaerobic granular sludge

To see if the general reported layered structure in anaerobic granules could be found in the granular sludge closer to the outside of the granule, a sample from the same WWTP was imaged at the edge of the granule, i.e. close to the surface of the granule towards the external environment (Sekiguchi et al., 1999; Satoh et al., 2007). The protein content was distributed unevenly in the mapped area. High absorbance height was found when going more towards the centre of the granule, as indicated by the white arrow. In contrast, polysaccharides were found to have a higher

absorbance height in closer towards the outer edge of the measured area. Lipid absorbance height had great similarity with the sialic acid, and protein distribution. With the lipid and sialic acid overlapping mostly in the region indicated with the white arrow in the sialic acid mapping. Sulfated compounds also was similar to the sialic acid and protein distribution but differed from sialic acids with a higher absorbance at the white arrow. The protein dominant region (indicated by the arrow), had a higher absorbance for the sulfated compounds than for the sialic acid compounds (figure 8).

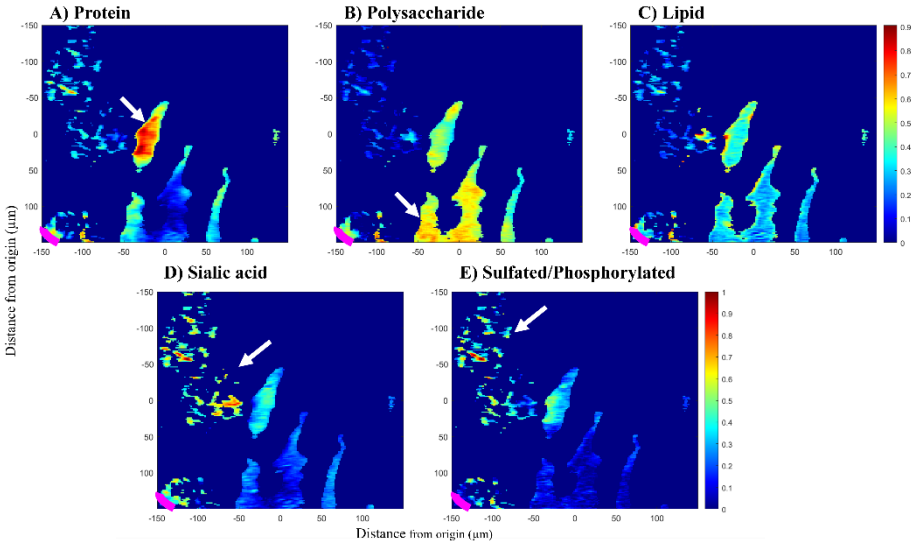


Figure 8. Scaled mapping of FT-IR functional groups absorbance band heights of sliced anaerobic granular sludge treating papermill wastewater imaged close to the edge of the granule. A) Height of the protein region absorbance band at 1630 cm-1. B) Height of the polysaccharide region absorbance band at 1030 cm-1. C) Height of the lipid region absorbance band at 2920 cm-1. D) Height of the sialic acid region absorbance band at 1726 cm-1. E) Height of the putative sulfate region absorbance band at 1206 cm-1. The origin in this picture is the centre of the ATR imaging crystal, and so also the centre of the imaged area. Regions that were filtered out due to weak signal are shown in dark blue. The white arrows indicate regions of high absorbance. The pink line indicates the outer edge of the granule.

3.5.1. Chemometric analysis on area on the outer edge of the granule.

Figure 9 shows the spectrum of the second, third and fourth PC loading spectra. Like the PCA performed on the centre part of the granule, the first principal component could explain the highest variability of the data (86.6%). The biggest variety was mostly caused by the presence of almost all functional groups. To image more specific functional groups, imaging was performed using the second, third and fourth PC (5.7%, 3.8%, 1.0%), respectively. Functional groups contributing to the second biggest variety of the data as seen in PC2 mostly polysaccharide in nature:

3300, 1084 and 1034 cm^{-1}. In PC2 protein absorbance bands were actually negatively contributing and composed of wavenumbers: 3060, 1676, 1580 and 1476 cm^{-1}. In contrast, PC3 was mainly protein dominated with absorbance bands at: 1620, 1530, 1450 and 1380 cm^{-1}. Other functional groups of interest where lipids found at absorbance bands: 2960 and 2870 cm^{-1}, and polysaccharides at: 1144 and 956 cm^{-1}. In PC4, protein absorbance bands were the main contributors in the wavenumbers: 1636 and 1540 cm^{-1}. Additionally, sialic acid absorbance bands at 1726 cm^{-1} contributed to PC4. Sulfate and polysaccharide were negatively contributing with the following wavenumbers: 1246, 1110, 948 and 854 cm^{-1}, respectively.

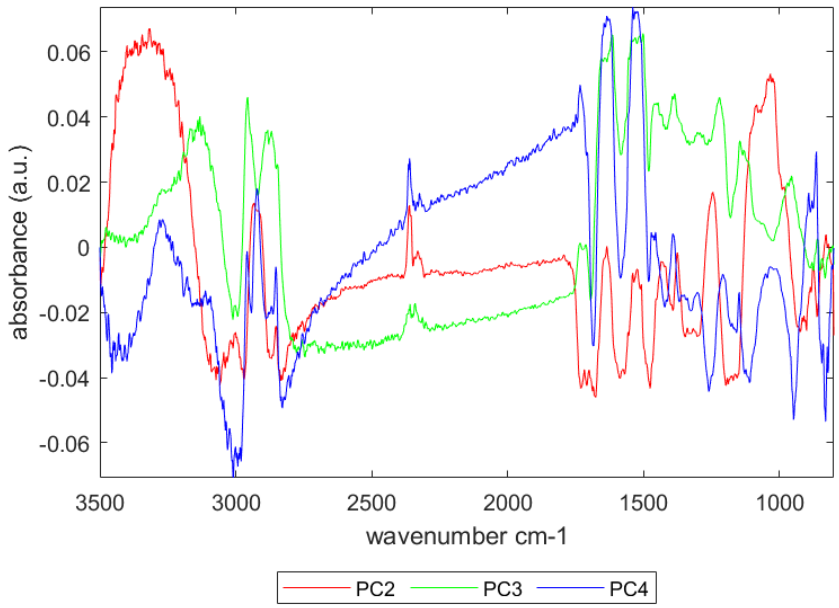


Figure 9. Principal components loading spectrum granular sludge slice at the outer edge.

The scoring for PC2, PC3 and PC4 of the spatially resolved spectra were mapped and are shown in figure 10. In the region close to the outer edge of the granule, high scores for PC2 (red) indicated a relatively higher presence of polysaccharides and sulfated polymers. Going toward the centre of the granule PC3 (green) and PC4 (blue) were more prevalent. High scores on PC3 and PC4 both indicated a protein dominant spectrum, but PC3 also contained absorbance for sulfated polymers and polysaccharide absorbance bands at 1154 and 956 cm^{-1}. Whereas, PC4 had a negative absorbance in the above mentioned polysaccharide bands, with a high absorbance at 854 cm^{-1}. The majority of the spectra scoring high on protein dominated PC loading spectra, were found towards the centre of the granule. Conversely, spectra scoring high on polysaccharides were found primarily near the edge of the granular sludge.

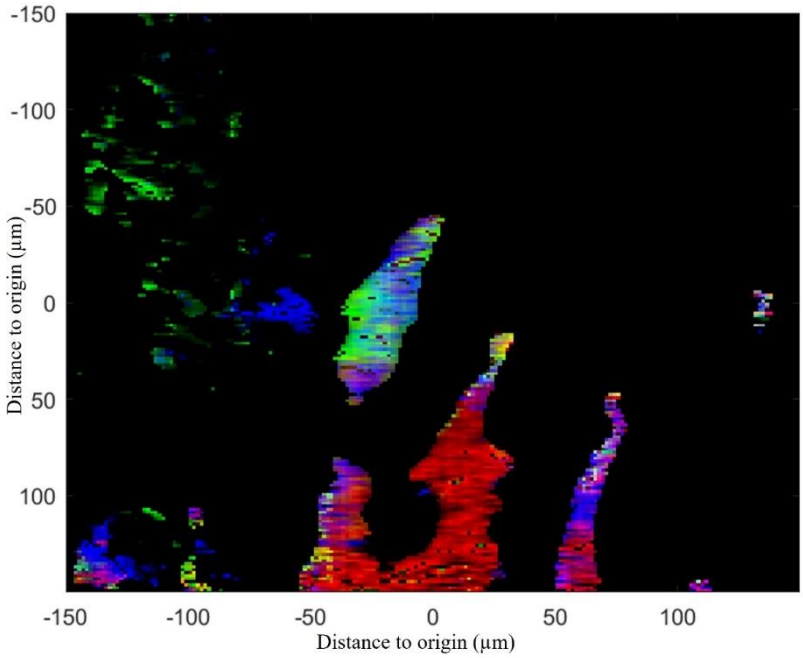


Figure 10. Mapping of the PC scores for PC2 (red), PC3 (green) and PC4 (blue) across the granule slice at the outer edge of the granule.

3.5.2. 2D-COS spectrum of the FT-IR spectra from the outer edge of the granule

The 2-D correlation spectrum for the FT-IR spectra is shown in figure 11. The most prominent functional groups visible in the data were sialic acid (1726 cm^{-1}), amide I & II (1630 & 1540 cm^{-1}), sulfated compounds (1234 cm^{-1}) and polysaccharides (1100 and 1024 cm^{-1}). The correlation spectrum on the area near the granule edge was similar to the correlation spectrum made from the data in the centre of the granule shown in figure 7.

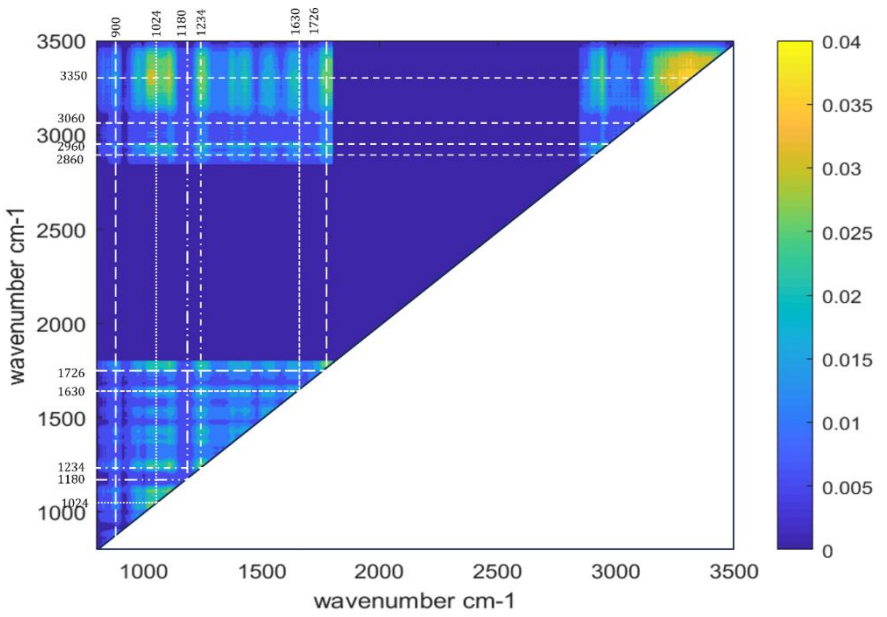


Figure 11. Mapping of the PC scores for PC2 (red), PC3 (green) and PC4 (blue) across the granule slice at the outer edge of the granule.

4. Discussion

4.1. The spatial structure of extracellular polymeric substances revealed by FT-IR microscopy

In this study, the possible application of FT-IR micro-spectroscopy for the visualization of extracellular polymer distribution across the granular sludge was explored. Anaerobic granular sludge has been described to have a layered structure. Fermentative bacteria grow on the outside, with syntrophic and hydrogenotrophic bacteria further in the middle and acetoclastic methanogens at the center of the granule (Satoh et al., 2007). In the center, the EPS is protein dominated (Ding et al., 2015; Forster & Quarmby, 1995). However, this layered structure is not always the case and clustered structures have also been reported (Gonzalez-Gil et al., 2001). It should be noted that the fermentative and syntropic layer in the layered structure can vary in size but also depends on the heterogeneity of the samples. In the current research, it was observed that there are regions with different EPS content which might follow the layered structure. While as far as it is close to the center of the granules, a more clustered structure can be distinguished. There seems to be no correlation between distance from the center and polymer composition. The granule sludge diameter was around 1.6 mm, estimated from the measured area size the covered area is 4.5 % of the total granular sludge slice. This mapped area would give an accurate representation of the sample if the EPS components were heterogeneously distributed. However, when looking at the data at the outer edge of the granules, a higher polysaccharide abundance at the surface and a higher protein abundance next to it towards the center of the granule is observed. The boundary

between the polysaccharide rich and the protein rich layers is about 150 μm from the granule surface, which is in accordance with other studies (Sekiguchi et al., 1999; Doloman et al., 2017). It is highly likely that the clustered EPS composition is caused by clustered communities of different microorganisms. The outside of the granular sludge seems to be dominated by a microbial community that produces a polysaccharide rich EPS, which could be attributed to fast-growing fermentative bacteria, which can easily obtain the substrate from the environment. Below the fermentative layer, the EPS composition did seem to be more randomly distributed and more clustered than layered.

4.2. Correlations in extracellular polymer composition unveiled by chemometric analysis

There were regions in the granular sludge slice that showed strong overlapping of functional groups in the EPS. This is partly caused by the convoluted absorbance spectra in biological structures (Marcott et al., 2014). Biological samples consist of complex polymers which are covalently linked proteins, polysaccharides, lipids and other molecules such as glycoproteins and glycolipids (Seviour et al., 2019). When covalently linked, they would co-localize in the spatial distribution. However, this is not the only reason why multiple absorbance bands would overlap in the spatial distribution map. Even a single polymer without covalent bonding has multiple functional groups that will generate multiple absorbance bands in varying degrees. By evaluating the increased or decreased absorbance in combinations with absorbance bands, it is possible to find features that are otherwise 'hidden' in the spectra (Jiang & Rieppo, 2006; Marcott et al., 2014; Ali et al., 2018). This is the basics of chemometrics, which focuses on exploratory data analysis that provides an overall view of the biological samples under study, allowing to catch possible similarities/dissimilarities among samples, to identify the presence of clusters or, in general, systematic trends, to discover which variables are relevant to describe the sample. The absorbance height figures of the single absorbance bands (figures 4 and 8) could visualize the presence of polysaccharides. However with the figures showing the PCA (figures 6 and 10), it could be seen which other functional groups would also be found present at the same place. Especially in figure 4, there are two regions that contain a high polysaccharide peak absorbance. In the PCA figure it is shown that one region scores high on PC2 and PC3, whereas the other region scored higher on PC4. By looking at the PCA scoring images (figures 6 and 10), it is observed that even though high polysaccharide regions were found at multiple locations, there can be different compositions in these regions. This distinction could be seen with PCA imaging. Therefore, chemometric techniques make it possible to look at multiple absorbance bands at the same time and process possible correlations between absorbance bands. With 2D correlation spectroscopy, it was observed that the functional groups of proteins and polysaccharides were positively correlated, indicating a possible presence of glycoproteins. In fact, these covalently linked polymers have been identified in many studies targeting granular sludge (Boleij et al., 2020; Chen et al., 2023; Felz et al., 2020). Sulfated compounds were also shown to be correlated with protein and polysaccharide functional groups, indicating a potential co-localization in the extracellular matrix. The generated 2D-

COS spectra of the different mapped areas were highly similar. However, differences could be observed in the positive correlations of sialic acids with polysaccharides. Figure 7 showed that sialic acids were positively correlated with polysaccharides and, in lesser extent, to proteins. Absorbance bands at 900 cm^{-1} and 1180 cm^{-1} were correlated to sialic acid absorbance bands. This correlation was not visible in figure 11, where a positive correlation was seen between sialic acid and absorbance bands at 900 and 1024 cm^{-1}. This could indicate that the sialic acids are linked to different sugar monomers or even completely different polymers. In Chen et al., (2023) it was shown that in the largest molecular weight samples, the EPS was more glycosylated and contained the most sialic acids. From the 2D-COS data this seems to be the case as well, whereas sulfated polymers are correlated to all different types of polymers.

4.3. Data analysis and limitations

The data acquired from the FT-IR micro-spectroscope is in the form of a hyperdimensional data cube. The FTIR images were acquired on the sliced granules across a 300 μm x 300 μm area, with a pixel size of 1.56 μm x 1.56 μm, which results in a total of 36,864 pixels. Within each pixel, the FT-IR absorbance spectrum was recorded across the 4000 – 600 cm-1 spectral range. Not all absorbance wavenumbers are relevant to the analysis of biological material. Given the interdependence of these wavenumbers and the selective presence of relevant ones in biological material, techniques like PCA and 2D-COS become invaluable for comprehensive data exploration. The inherent complexity of biological tissue FT-IR data necessitates the incorporation of statistical methods to improve understanding of the FT-IR data (Marcott et al., 2014; Ali et al., 2018). Through the application of chemometric data analysis, it is possible to discern correlations among functional groups and visualize the distribution of multiple functional groups across the granule. This approach enhances the ability to identify and understand the relationships among different chemical constituents within the studied samples.

The utility of techniques such as PCA and 2D-COS extends beyond only FT-IR microscopy, as it can be applied for any spectroscopic measurement method such as Raman microscopy, Imaging Mass Spectrometry and Magnetic Resonance Imaging, among others (Gowen et al., 2015). Their versatility lies in their capacity to unravel complex datasets, identify correlations, and provide insightful visualizations not confined to a single analytical domain. This underscores their broader importance as powerful tools for comprehensive data exploration and interpretation across various scientific disciplines.

The FT-IR micro-spectroscope proved highly sensitive to pressure fluctuations affecting the sample's surface area, necessitating vigilant maintenance of proper contact between the sample and the ATR crystal during measurements. In the context of similar biological material compositions, differentiation between regions inside mapped areas is intricate, and complicated by the convolution of absorbance bands from overlapping functional groups. The penetration depth (up to 2 μm), as well as the spatial resolution of 1.56 μm in this FT-IR micro-spectroscope poses challenges in attributing specific observations to individual microorganisms, making

e.g. microbial identification less straightforward compared to, for instance, laser microscopy (Gowen et al., 2015). It is also due to these limitations, that distinguishing between intracellular and extracellular components becomes challenging. Validation of observed polymers as part of the extracellular matrix, rather than the intracellular space, requires complementary techniques. In the future, it would be beneficial to perform validation studies using molecular imaging techniques of microbial composition i.e. FISH or with chemical or biological (lectin) staining methods. Nonetheless, this non-destructive imaging method could be a powerful imaging method for imaging the EPS distribution across the granular sludge.

Conclusion

In this study, FT-IR micro-spectroscopy was used to explore and visualize the EPS spatial distribution inside anaerobicgranular sludge. FT-IR micro-spectroscopy is an untargeted imaging method that can measure a broad range of polymers simultaneously. With the single-point absorbance it was possible to create an absorbance height distribution for multiple different functional groups corresponding to proteins, polysaccharides, lipids, sialic acid and phosphorylated/sulfated polymers. Due to the complexity of the extracellular matrix, chemometric techniques were implemented to help elucidate the EPS composition. Principal component analysis made it possible to visualize multiple functional groups at the same time, which helped in distinguishing between regions that high absorbance height in for similar functional groups. Two-dimensional correlation spectroscopy allowed for the identification of correlated functional groups. In this study we show that FT-IR micro-spectroscopy is a powerful complementary technique for EPS analysis and visualization. The combination of this non-destructive imaging method with other molecular imaging techniques would provide valuable insights in the role of specific polymers inside the extracellular matrix in anaerobic granular sludge

References

Ali, M. H., Rakib, F., Al-Saad, K., Al-Saady, R., Lyng, F. M., & Goormaghtigh, E. (2018). A simple model for cell type recognition using 2D-correlation analysis of FTIR images from breast cancer tissue. *Journal of Molecular Structure*, *1163*, 472-479.

Boleij, M., Kleikamp, H., Pabst, M., Neu, T. R., van Loosdrecht, M. C., & Lin, Y. (2020). Decorating the anammox house: sialic acids and sulfated glycosaminoglycans in the extracellular polymeric substances of anammox granular sludge. *Environmental science & technology*, *54*(8), 5218-5226.

Bourven, I., Bachellerie, G., Costa, G., & Guibaud, G. (2015). Evidence of glycoproteins and sulphated proteoglycan-like presence in extracellular polymeric substance from anaerobic granular sludge. *Environmental technology*, *36*(19), 2428-2435.

Bou-Sarkis, A., Pagliaccia, B., Ric, A., Derlon, N., Paul, E., Bessiere, Y., & Girbal-Neuhauser, E. (2022). Effects of alkaline solvents and heating temperatures on the solubilization and degradation of gel-forming Extracellular Polymeric Substances (EPS) extracted from aerobic granular sludge. *Biochemical Engineering Journal*, *185*, 108500.

Brézillon, S., Untereiner, V., Lovergne, L., Tadeo, I., Noguera, R., Maquart, F. X., ... & Sockalingum, G. D. (2014). Glycosaminoglycan profiling in different cell types using infrared spectroscopy and imaging. *Analytical and bioanalytical chemistry*, *406*, 5795-5803.

Chen, L. M., de Bruin, S., Pronk, M., Sousa, D. Z., van Loosdrecht, M. C., & Lin, Y. (2023). Sialylation and Sulfation of Anionic Glycoconjugates Are Common in the Extracellular Polymeric Substances of Both Aerobic and Anaerobic Granular Sludges. *Environmental Science & Technology*, *57*(35), 13217-13225.

Chirman, D., & Pleshko, N. (2021). Characterization of bacterial biofilm infections with Fourier transform infrared spectroscopy: A review. *Applied Spectroscopy Reviews*, *56*(8-10), 673-701.

Costa, O. Y., Raaijmakers, J. M., & Kuramae, E. E. (2018). Microbial extracellular polymeric substances: ecological function and impact on soil aggregation. *Frontiers in microbiology*, *9*, 1636.

De Bruin, S., Vasquez-Cardenas, D., Sarbu, S. M., Meysman, F. J. R., Sousa, D. Z., van Loosdrecht, M. C. M., & Lin, Y. (2022). Sulfated glycosaminoglycan-like polymers are present in an acidophilic biofilm from a sulfidic cave. *Science of the Total Environment*, *829*, 154472.

de Graaff, D. R., Felz, S., Neu, T. R., Pronk, M., van Loosdrecht, M. C., & Lin, Y. (2019). Sialic acids in the extracellular polymeric substances of seawater-adapted aerobic granular sludge. *Water research*, *155*, 343-351.

Ding, Z., Bourven, I., Guibaud, G., van Hullebusch, E. D., Panico, A., Pirozzi, F., & Esposito, G. (2015). Role of extracellular polymeric substances (EPS) production in bioaggregation: application to wastewater treatment. Applied microbiology and biotechnology, 99, 9883-9905.

Doloman, A., de Bruin, S., van Loosdrecht, M. C., Sousa, D. Z., & Lin, Y. (2024). Coupling extracellular glycan composition with metagenomic data in papermill and brewery anaerobic granular sludges. Water Research, 121240.

Doloman, A., Varghese, H., Miller, C. D., & Flann, N. S. (2017). Modeling de novo granulation of anaerobic sludge. BMC systems biology, 11, 1-12.

Dubé, C. D., & Guiot, S. R. (2019). Characterization of the protein fraction of the extracellular polymeric substances of three anaerobic granular sludges. *AMB Express*, *9*, 1-11.

Felz, S., Al-Zuhairy, S., Aarstad, O. A., van Loosdrecht, M. C., & Lin, Y. M. (2016). Extraction of structural extracellular polymeric substances from aerobic granular sludge. *JoVE (Journal of Visualized Experiments)*, (115), e54534.

Felz, S., Neu, T. R., van Loosdrecht, M. C., & Lin, Y. (2020). Aerobic granular sludge contains Hyaluronic acid-like and sulfated glycosaminoglycans-like polymers. *Water research*, *169*, 115291.

Felz, S., Vermeulen, P., van Loosdrecht, M. C., & Lin, Y. M. (2019). Chemical characterization methods for the analysis of structural extracellular polymeric substances (EPS). *Water Research*, *157*, 201-208.

Flemming, H. C., & Wingender, J. (2010). The biofilm matrix. *Nature reviews microbiology*, *8*(9), 623-633.

Forster, C. F., & Quarmby, J. (1995). The physical characteristics of anaerobic granular sludges in relation to their internal architecture. Antonie van Leeuwenhoek, 67, 103-110.

Gagliano, M. C., Neu, T. R., Kuhlicke, U., Sudmalis, D., Temmink, H., & Plugge, C. M. (2018). EPS glycoconjugate profiles shift as adaptive response in anaerobic microbial granulation at high salinity. *Frontiers in Microbiology*, *9*, 1423.

Gonzalez-Gil, G., Lens, P. N. L., Van Aelst, A., Van As, H., Versprille, A. I., & Lettinga, G. (2001). Cluster structure of anaerobic aggregates of an expanded granular sludge bed reactor. Applied and Environmental Microbiology, 67(8), 3683-3692.

Gonzalez-Gil, G., Thomas, L., Emwas, A. H., Lens, P. N., & Saikaly, P. E. (2015). NMR and MALDI-TOF MS based characterization of exopolysaccharides in anaerobic microbial aggregates from full-scale reactors. *Scientific Reports*, *5*(1), 14316.

Gowen, A. A., Feng, Y., Gaston, E., & Valdramidis, V. (2015). Recent applications of hyperspectral imaging in microbiology. *Talanta*, *137*, 43-54.

Jiang, E. Y., & Rieppo, J. (2006). Enhancing FTIR imaging capabilities with two-dimensional correlation spectroscopy (2DCOS): A study of concentration gradients of collagen and proteoglycans in human patellar cartilage. *Journal of molecular structure*, *799*(1-3), 196-203.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150202.

Lasch, P., & Noda, I. (2019). Two-dimensional correlation spectroscopy (2D-COS) for analysis of spatially resolved vibrational spectra. *Applied Spectroscopy*, *73*(4), 359-379.

Li, P. N., Herrmann, J., Tolar, B. B., Poitevin, F., Ramdasi, R., Bargar, J. R., ... & van den Bedem, H. (2018). Nutrient transport suggests an evolutionary basis for charged archaeal surface layer proteins. *The ISME journal*, *12*(10), 2389-2402.

Lotti, T., Carretti, E., Berti, D., Montis, C., Del Buffa, S., Lubello, C., ... & Malpei, F. (2019). Hydrogels formed by anammox extracellular polymeric substances: Structural and mechanical insights. *Scientific reports*, *9*(1), 11633.

Marcott, C., Lo, M., Hu, Q., Kjoller, K., Boskey, A., & Noda, I. (2014). Using 2D correlation analysis to enhance spectral information available from highly spatially resolved AFM-IR spectra. *Journal of molecular structure*, *1069*, 284-289.

Nallala J, Jeynes C, Saunders S, Smart N, Lloyd G, Riley L, Salmon D, Stone N. Characterization of colorectal mucus using infrared spectroscopy: a potential target for bowel cancer screening and diagnosis. Lab Invest. 2020 Aug;100(8):1102-1110. doi: 10.1038/s41374-020-0418-3. Epub 2020 Mar 20. PMID: 32203151; PMCID: PMC7374084.

Seviour, T., Derlon, N., Dueholm, M. S., Flemming, H. C., Girbal-Neuhauser, E., Horn, H., ... & Lin, Y. (2019). Extracellular polymeric substances of biofilms: Suffering from an identity crisis. *Water research*, *151*, 1-7.

Neu, T. R., & Kuhlicke, U. (2022). Matrix glycoconjugate characterization in multispecies biofilms and bioaggregates from the environment by means of fluorescently-labeled lectins. *Frontiers in Microbiology*, *13*, 940280.

Noda, I., Dowrey, A. E., Marcott, C., Story, G. M., & Ozaki, Y. (2000). Generalized two-dimensional correlation spectroscopy. *Applied Spectroscopy*, *54*(7), 236A-248A.

Pinel, I. S., Kim, L. H., Proença Borges, V. R., Farhat, N. M., Witkamp, G. J., van Loosdrecht, M. C., & Vrouwenvelder, J. S. (2020). Effect of phosphate availability on biofilm formation in cooling towers. *Biofouling*, *36*(7), 800-815.

Satoh, H., Miura, Y., Tsushima, I., & Okabe, S. (2007). Layered structure of bacterial and archaeal communities and their in situ activities in anaerobic granules. Applied and environmental microbiology, 73(22), 7300-7307.

Sekiguchi, Y., Kamagata, Y., Nakamura, K., Ohashi, A., & Harada, H. (1999). Fluorescence in situ hybridization using 16S rRNA-targeted oligonucleotides reveals localization of methanogens and selected uncultured bacteria in mesophilic and thermophilic sludge granules. Applied and environmental microbiology, 65(3), 1280-1288.

Talari, A. C. S., Martinez, M. A. G., Movasaghi, Z., Rehman, S., & Rehman, I. U. (2017). Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews*, *52*(5), 456-506.

Van Lier, J. B., Van der Zee, F. P., Frijters, C. T. M. J., & Ersahin, M. E. (2015). Celebrating 40 years anaerobic sludge bed reactors for industrial wastewater treatment. *Reviews in Environmental Science and Bio/Technology*, *14*, 681-702.

Supplemental information S1

MATLAB script for FT-IR imaging

```
clearvars

% Load fsm data obtained from PerkinElmer frontier 400
micro spectroscope
% Data consists of an X Y Z block where X and Y are
location parameters and Z is the absorbance spectrum on
each X,Y location.
% Unfold the Z data block from [X,Y] to [X.Y,1] to ease
processing.
[data, xAxis, yAxis, zAxis, ~] =
fsmload('Path/To/Filename/Filename.fsm');
spectrum =
zeros(length(xAxis)*length(yAxis),length(zAxis));
n = 1;
for i = 1:length(xAxis)
    for j = 1:length(yAxis)
        spectrum(n,:) = data(i,j,:);
        n = n + 1;
    end
end

% Change transmittance to absorbance
abs_sp = 2 - log10(spectrum);
clear data spectrum

% Create a subset for region of interest (3500 to 800 cm-1)
sp = abs_sp(:,250:end-55);
zAxis_n = zAxis(250:end-55);
clear zAxis abs_sp n

% Smoothing (Savitsky Golay)
sp = smoothdata(sp(:,:),'sgolay',10);

% Simple baseline correction based on first order to
correct baseline drift
a = (sp(:,1) - sp(:,end)) ./ (3500 - 800);
corr = a .* zAxis_n;
bl_sp = sp(:,:) - corr(:,:);
bl_sp = bl_sp - bl_sp(:,1);
bl_sp = abs(bl_sp);
bl_sp = normalize(bl_sp,1,'range',[0 1]);
clear a corr s

% Filter out bad data points
a = mean(bl_sp,2);
BL_SP = a > 0.3;
bl_sp = bl_sp .* BL_SP;
```

```
% Plot average absorbance data
b = reshape(a,length(yAxis),length(xAxis));
b = b';
figure
image(xAxis,yAxis,b,'CDataMapping','scaled');
truesize
clear a

% Principal component analysis
[coeff,score,latent,tsquared,mu,explained] = pca(bl_sp);
% Plot the coefficient spectra/loading spectra of the first
5 PCs
figure
plot(zAxis_n,coeff(:,1),'r',zAxis_n,coeff(:,2),'g',zAxis_n,
coeff(:,3),'b',zAxis_n,coeff(:,4),'k',zAxis_n,coeff(:,5))
set(gca,'XDir','reverse');
legend('PC1','PC2','PC3','PC4','PC5');

% Imaging absorbance of single absorbance peaks
p1 = 1630; p2 = 1030; p3 = 2920; p4 = 1726; p5 = 1206;
C2 = zeros(192,192,5);
abs_p = zeros(192,192);
peaks = [ p1; p2; p3 ;p4;p5];
for p = 1:5
    for m = 1:length(xAxis)
        abs_p(m,:) =  bl_sp((m-
1)*length(yAxis)+1:m*length(yAxis),find(zAxis_n==peaks(p)))
;
    end
    C2(:,:,p) = abs_p;
end

% Imaging three functional groups at the same time
C3 = normalize(C2,1,'range');
figure;
image(xAxis1,yAxis1,data1);
hold on
image(xAxis,yAxis,C3(:,:,[1 4 5]));
title('polymers');
hold off

% Imaging PCA
pc_score = zeros(length(xAxis),length(yAxis));
part1 = score(1:length(xAxis)*length(yAxis),:);
C = zeros(length(xAxis),length(yAxis),3);
Cn = zeros(length(xAxis),length(yAxis),3);
color = [1,2,3]; % PC's chosen
for p = 1:3
    for m = 1:length(xAxis)
```

```
        pc_score(m,:) = part1((m-
1)*length(yAxis)+1:m*length(yAxis),color(p));
    end
    C(:,:,p) = pc_score;
end
C(C<0)=0;
% Normalize
Cn = normalize(C,1,'range');
b(b<0.3)=0; b(b>0.3)=1;
% Create image
figure;
image(xAxis1,yAxis1,data1);
hold on
image(xAxis,yAxis,Cn.*b);
hold off

% Two-D COS
A_av = mean(bl_sp,1);
A_vk = bl_sp - A_av;
A_vk(:,330:850)=0;
dot = cov(A_vk);
diago = diag(dot);
figure;
plot(zAxis_n,diago)
figure;
contourf(zAxis_n,zAxis_n,dot,'EdgeColor','none');
```

Supplemental information

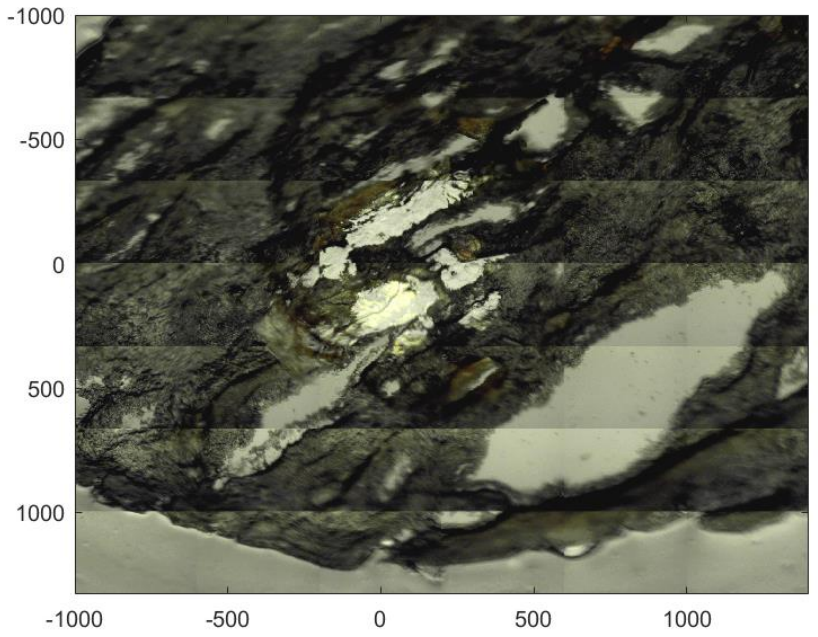


Figure S1. Sliced anaerobic granule used for FT-IR micro-spectroscopy. X and Y axis display distance from origin in μm.

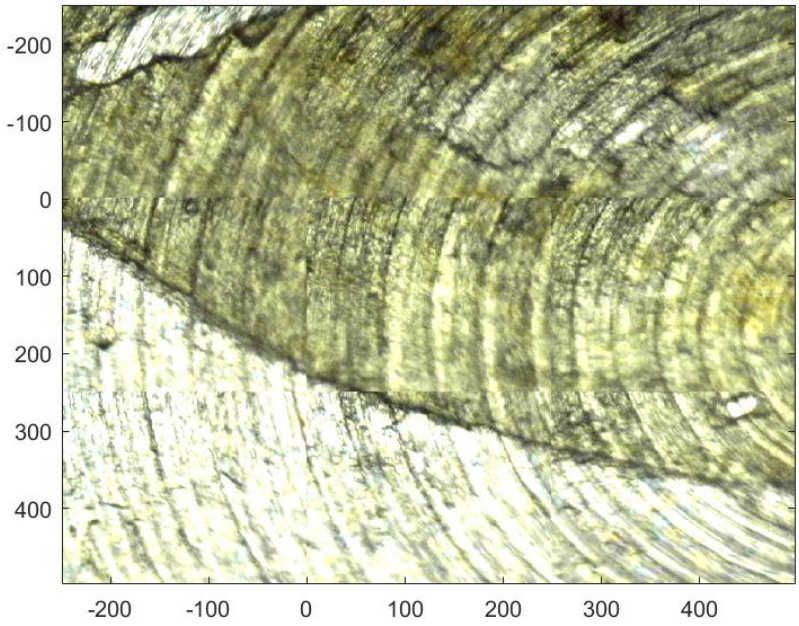


Figure S2. Sliced anaerobic granule used for FT-IR micro-spectroscopy. The background seen here is a metal wafer, hence the rings. X and Y axis display distance from origin in μm.

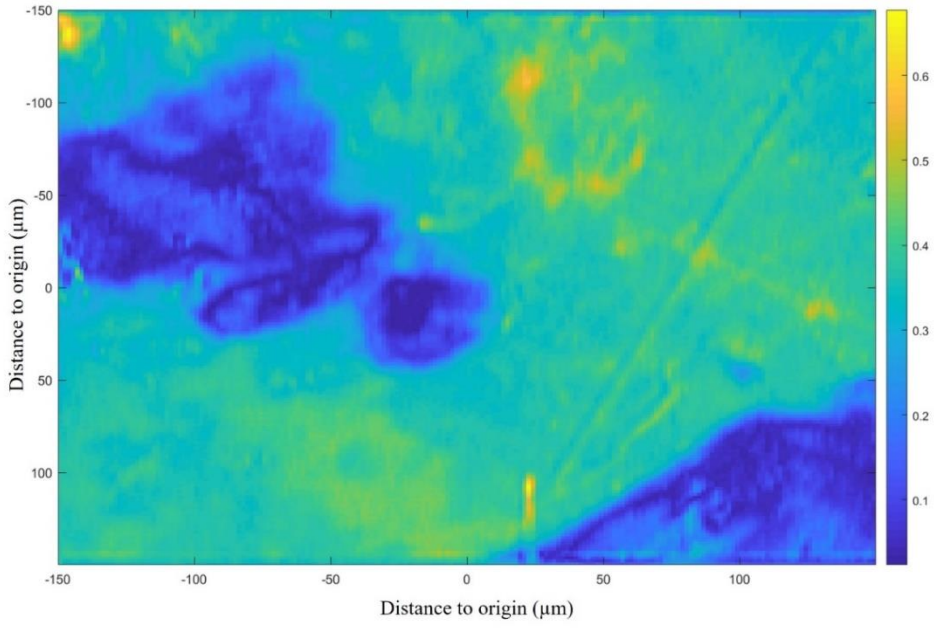


Figure S3. Average absorbance height (a.u.) of the FT-IR spectrum measured for granule (S1) for a 300 by 300 μm area.

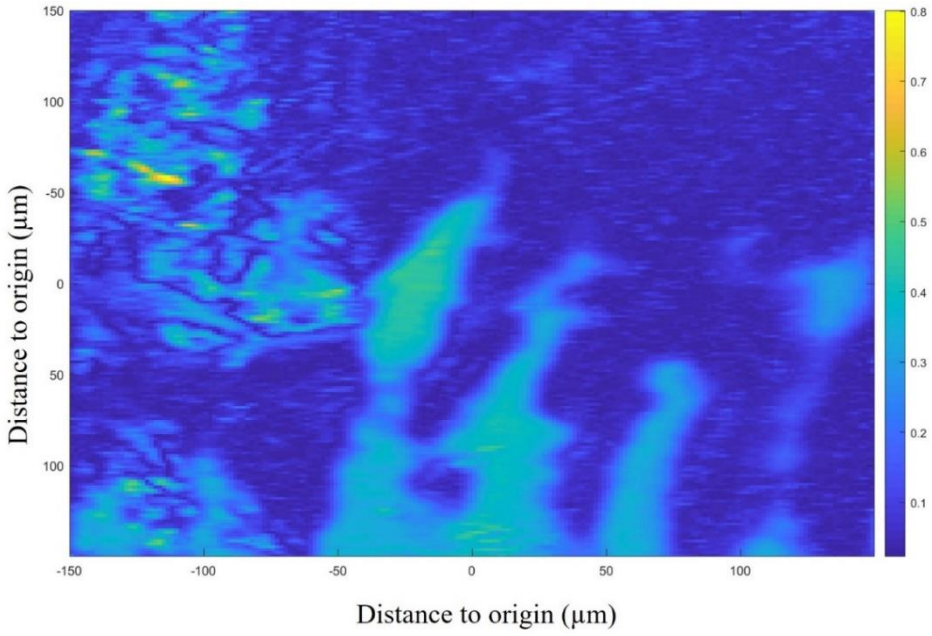


Figure S4. Average absorbance height (a.u.) of the FT-IR spectrum measured for granule (S2) for a 300 by 300 μm area.

Chapter 6:
Outlook

In spite of industrial wastewater treatment with anaerobic granular sludge reactors having been an established process for over more than 40 years, there are still many unknowns in the composition and role of extracellular polymeric substances (EPS) in these systems. Driven by the desire understanding and knowledge expansion in this domain, investigations described on this thesis have delved into different aspects of structure and function of microbes and their matrixes in anaerobic granular sludge. When the anaerobic granular sludge is conceptualized as a microbial city, microorganisms can be envisioned as its inhabitants, macro-structures within the extracellular matrix as the city's buildings, and EPS components as the fundamental building blocks of the structures. Throughout the chapters in this thesis, new methodologies have been developed to explore the building blocks and touch upon the inhabitants within the microbial city. Fundamental questions regarding: what type of EPS can be found, how the microbial community influences EPS composition, and how the EPS composition varies at different location in the sludge were addressed (see figure 6.1 for the conceptualized representation).

Figure 6.1. Schematic representation of the questions (partly) answered in this thesis. From left to right: What building blocks (EPS) can be found in the extracellular matrix? How do the inhabitants of the granules influence the EPS composition? How does the lay-out/location influence the EPS composition?

6.1. EPS building blocks: extraction and isolation of acidic polymers

The extraction of EPS is a recurring subject in the field, and previous PhD researchers have highlighted the importance of solubilizing the extracellular matrix as a prerequisite step. Often, harsh extraction methods were deemed necessary to achieve this goal (Felz 2019; Boleij 2020; Tomas Martinez 2023). Coming back to the analogy of anaerobic granular sludge to the microbial city, this harsh extraction is akin to razing all the buildings in the microbial city and collect as much debris as possible. Within the debris, there are different building blocks. In this thesis, focus has been put on establishing methods for the isolation and characterization of specific building blocks. These building blocks are acidic polymers, e.g. glycoconjugates with sulfated glycosaminoglycan-like polymers and sialic acids. In **Chapter 2** by studying an acidophilic biofilm collected from Sulfur Cave in Puturosu Mountain (Romania), an environment that is largely inaccessible to contamination, it is shown that sulfated glycosaminoglycan-like polymers are produced inside the biofilm by the microorganisms instead of accumulating from external sources. The presence of these polymers in various other biofilms imply that these are also produced internally (Felz 2019; Boleij 2020). This shows that these polymers probably have an universal role in across biofilms. In **Chapter 3** a size-exclusion separation approach was highlighted for isolating polymers of different molecular weight. Often when EPS with different molecular weight are separated based on size, some EPS fractions are too big to pass through the gel, making large molecular weight fractions underrepresented in the data. In this approach, we addressed this concern and showed that 16 - 27 % of the EPS had an apparent molecular weight of >5,500 kDa. The high molecular weight fraction had the highest concentration of sialic acids. Interestingly, sulfated polymers were distributed over all the different molecular weight fractions. This result revealed that sulfated and sialylated polymers probably have different functions in the extracellular matrix.

In this thesis, solubilizing a large fraction of EPS with harsh extractions is still the focus. This approach is good for exploratory analysis but challenges structural analysis. This challenge comes from interference from similar functional groups which convolute any analytical methods. Size exclusion chromatography was used in this thesis to separate different apparent molecular weight polymers and thereby provide more enriched samples. The term apparent molecular weight was used as determining the molecular weight is challenging because of for instance interactions of polymer-polymer or column-polymer nature. only be approximated by our observations. These interactions between the EPS and column influence the

separation, which is why the fractions might have a broad apparent molecular weight range.

6.2. Microbial influence on EPS composition

The microbial community determines the genetic potential for the production of EPS and therefore must also influence the EPS composition. It was shown in **Chapter 4** that, for anaerobic granular sludge, the genera *Anaerolinea* and members of the *Methanobacter* class contained the highest genetic potential for producing nucleotide sugar monomers of the dominant sugar monomers present in EPS. Through the knowledge of the sugar monomer composition in the EPS, it is possible to guide the search for EPS synthesis genes related to the most dominant sugars present in the EPS. This resulted in an approach for finding EPS synthesis genes related to real-time EPS components, rather than only relying on the knowledge of possible EPS components reported in literature. **Chapter 5** showed how the EPS composition differed in the extracellular matrix structure across the anaerobic granule. Anaerobic granular sludge is described to have a layered structure, based on the distribution of microorganisms performing different steps of the anaerobic digestion cascade. The ability to see a change in the EPS composition at approximately 150 μm from the granule edge coincided with the size of the acidogenic microorganism layer found in previous studies. Preliminary findings (not included in previous chapters), using protein and polysaccharide colorimetric methods as described in **Chapter 3**, suggest that granule size influences the Protein/Polysaccharide (PN/PS) ratio, with larger granules (>1 mm), medium granules (0.3-1 mm), and small granules (<0.3 mm) exhibiting PN/PS ratios of 8.8, 5.5, and 3.7, respectively. Amplicon 16S rRNA sequencing revealed differences in microbial communities across granule sizes, illustrated through a heatmap showcasing correlations between microbial families (figure 6.3), i.e., larger granules showed higher relative abundance of bacterial families (*Geobacteriaceae* and *Anaerolineaceae*) compared to methanogenic families (*Methanosaetaceae* and *Methanobacteriaceae*), indicating a potential relationship between granule size, microbial composition, and EPS characteristics.

However, how these microorganisms influence the EPS composition is still largely unknown. Currently, only limited EPS operons have been described in literature, with many EPS synthesis pathways being largely unknown. It must also be noted that genetic potential does not equal to production, so the genetic ability of one microorganism to produce EPS does not mean this microorganism is definitely the EPS producer.

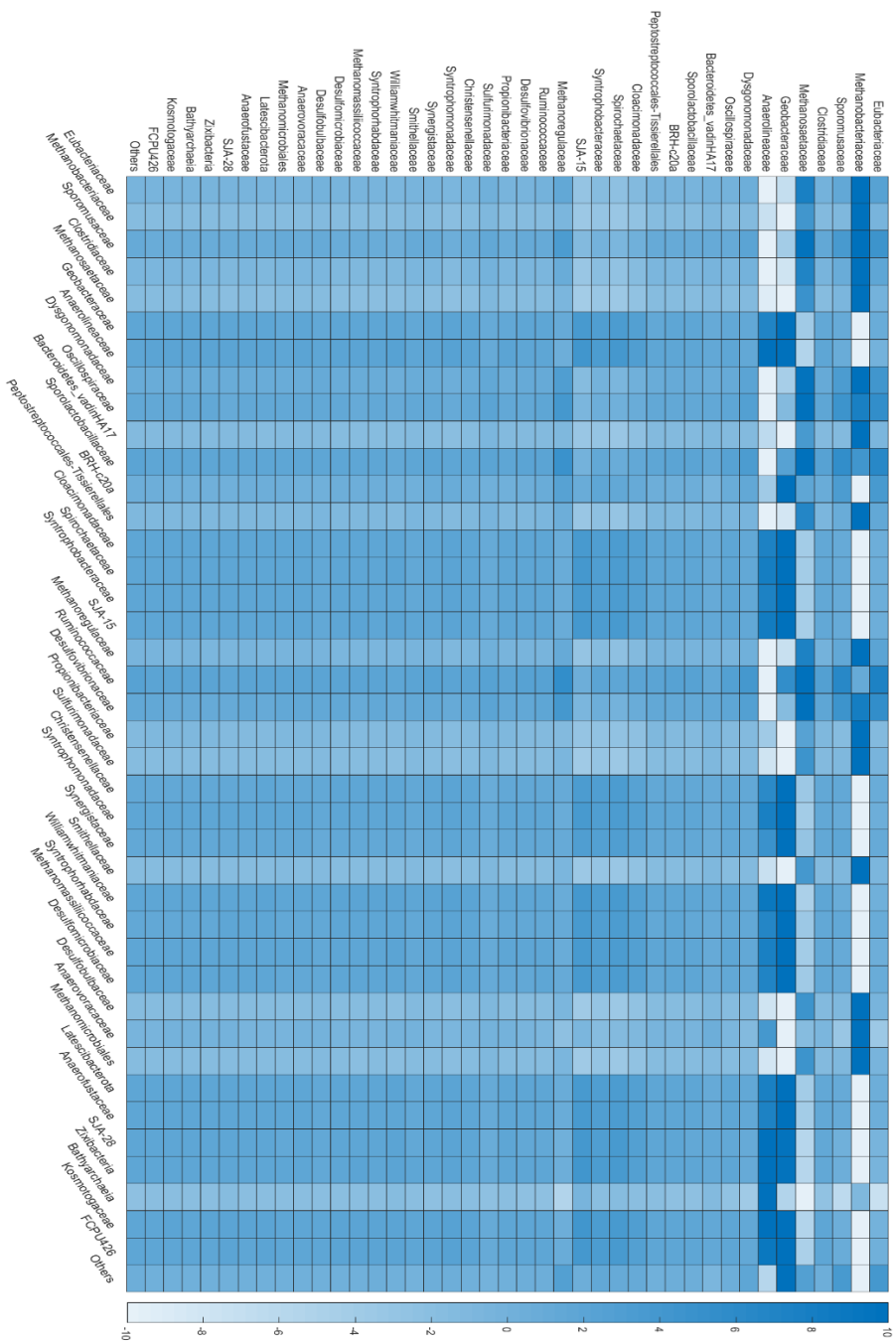


Figure 6.2. Amplicon sequencing 16S rRNA relative abundance heatmap displaying the microbial family covariance across the different granule sizes

6.3. EPS composition and localization across the granule extracellular matrix

Chapter 5 showed how the EPS composition differed in the extracellular matrix structure across the anaerobic granule through FTIR microscopy. The methods demonstrated are not only useful to visualize anaerobic granular sludge but are promising for visualizing other types of sludges and biofilms in general as well. FT-IR micro-spectroscopy detects a diverse array of polymers simultaneously. Consequently, it facilitated the determination of spatial distribution patterns of various polymers within the EPS matrix across the sludge. Principal component analysis made it possible to visualize which functional groups caused the biggest variance in the data, which helped in distinguishing what functional groups were important in the extracellular matrix. Two-dimensional correlation spectroscopy allowed for identifying how functional groups change across the granular sludge slice and could show which groups are correlated with each other. Both these two methods helped in decluttering the FT-IR spectral data and aid in annotating specific polymers to specific areas in the granule.

Raman micro-spectroscopy is an imaging technique that is complementary to FT-IR imaging. In preliminary work performed at Aalborg University with Raman micro-spectroscopy, the feasibility of using Raman micro-spectroscopy for imaging the EPS composition was tested. The set-up allowed for the integration of Raman micro-spectroscopy and molecular imaging (Petriglieri et al., 2021). A proof of concept was performed with a pure culture *Pseudomonas aeruginosa* biofilm. The ability to differentiate between the microorganisms and the extracellular matrix by Raman micro-spectroscopy was explored. To do this, the biofilm was DAPI stained and subsequently Raman spectra were recorded from DAPI positive and negative pixels (figure 6.3). It was possible to determine differences in the average spectra for DAPI positive and negative points (figure 6.4). The DAPI-stained mapping points specifically had a higher absorbance at 995 and 1145 cm^{-1}, which indicated a relatively higher protein content compared to the non-stained points. The non-stained points had higher absorbances at 726, 822 and 917 cm^{-1}, indicating polysaccharide C-C stretching. The advantage of Raman compared to FT-IR micro-spectroscopy is that location selection for the absorbance measurements is more accurate, and the spatial resolution is higher. The proof of concept study showed that Raman micro-spectroscopy could be an interesting technique for analysing the extracellular matrix around specifically stained microorganisms.

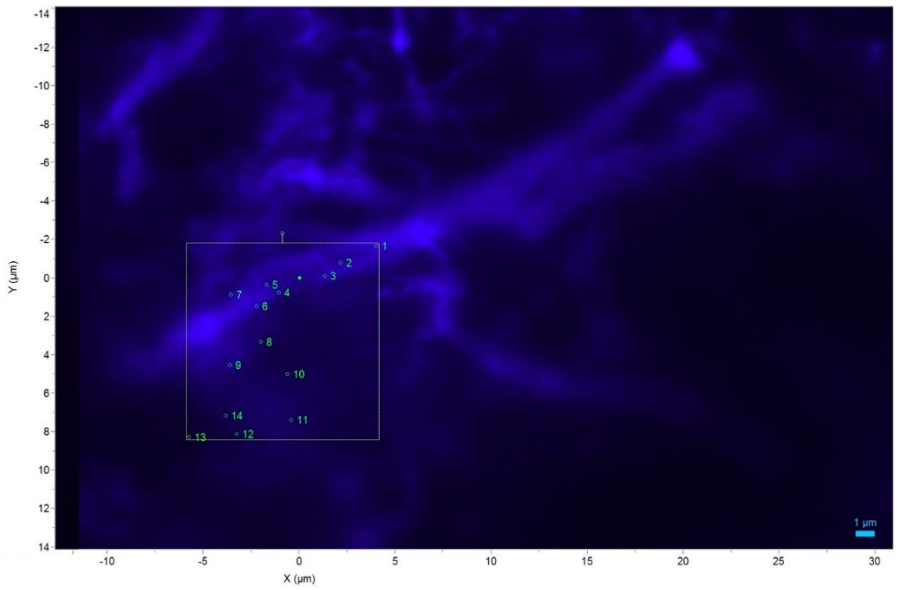


Figure 6.3. Raman mapping of DAPI stained filaments and extracellular matrix. Numbers (1-7) are DAPI-positive locations and (8-14) are DAPI negative locations.

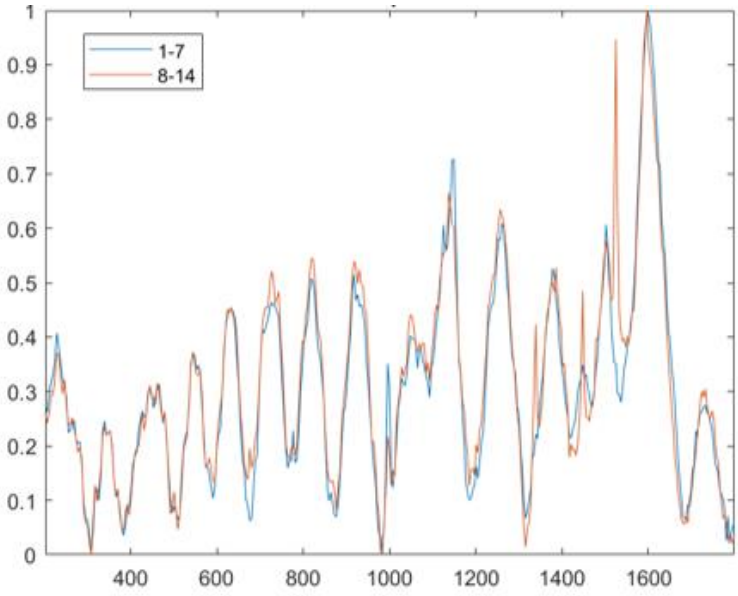


Figure 6.4. Average Raman spectra between mappings on DAPI positive locations (1-7) and DAPI negative locations (8-14).

It is noted that FT-IR imaging provides a spatial distribution map of a broad range of polymers present in the extracellular matrix. However, due to overlapping of functional groups present in all biological materials, many of the results can be convoluted. This leads to a more generalized representation of the extracellular matrix, as only general polymer groups can be distinguished. This might be a limiting factor when distinguishing specific polymers such as bacterial sialic acids (e.g. pseudaminic acid and legionaminic acid) from their sialic acid analogues. In addition, the heterogeneous nature of the biofilm makes it difficult to generate a clear indication from the patterns found in the data. Moreover, differences in the sample thickness can interfere with the absorbance spectrum. Proper sample handling is therefore important for successful imaging.

6.4. Recommendations for future work

It is worth pointing out that, selective extraction and isolation methods (e.g., affinity chromatography) reduce the complexity of the recovered EPS, facilitating a more detailed analysis. Extraction with sequentially harsher treatment results in a enriched EPS fractions with simplified compositions. Another way to simplify the complexity of EPS mixture is by increasing the separation resolution of size-exclusion chromatography, as conducted in **Chapter 3**. Optimizations may involve using different eluents, varying polymer concentrations, and adjusting polymer loading volumes together with a thorough research on the different fractions collected. In addition to enhancing our comprehension of acidic polymers, elucidating the structure of the most prevalent polysaccharides can further guide the analysis on genes related to EPS production. The "EPS analytical data to genomic search approach" as demonstrated in **Chapter 4**, would be more powerful when complete polymer structures are used for filtering EPS synthesis genes.

The enriched EPS fractions facilitate the determination of polysaccharide structure using advanced analytical methods. Linkage analysis via GC-MS identifies the hydroxyl groups attached to sugar groups, and can thus determine the polysaccharide structure (Black et al., 2021). Furthermore, NMR analysis, such as NOESY, can establish the sequence of sugar monomers in polysaccharides (Wang et al., 2021). This structural information can help the exploration of potential biosynthesis clusters through metagenomics with gene proximity analysis (Dueholm et al., 2023). However, it is recommended to additionally measure actual production through transcriptomics or proteomics in addition to assessing production potential.

Regarding visualization, with spectral imaging like e.g. FT-IR and Raman micro-spectroscopy, a broad range of polymers can be visualized to track the spatial distribution of different polymers, but the bottleneck is how to make annotation on the complex data. Validation of the data by localizing specific polymers through chemical staining or fluorescence lectin staining analysis would help to narrow down the possibility and understand the extracellular matrix structure better. How microorganisms contribute to the matrix structure could then be further analysed with the combination of spectral imaging and molecular imaging methods like e.g. Fluorescence In-Situ Hybridization (Seviour et al., 2019).

Lastly, the effect of the granule samples origin and physical parameters should be further investigated. Currently, the influence of granule size on the EPS composition has not been well taking into account in EPS research. In our preliminary study, it was shown that the granule size does influence the microbial community and the EPS composition. Separating granular sludge based on size might therefore be a useful way to enrich for specific polymers. Performing the spectral imaging on granular sludge from lab-scale reactors would be a great comparison in understanding the granular sludge and EPS composition, because of the more well-defined media and operational conditions.

References

Black, I., Heiss, C., Carlson, R.W., Azadi, P. (2021). Linkage Analysis of Oligosaccharides and Polysaccharides: A Tutorial. In: Delobel, A. (eds) Mass Spectrometry of Glycoproteins. Methods in Molecular Biology, vol 2271. Humana, New York, NY. https://doi.org/10.1007/978-1-0716-1241-5_18

Boleij, M. (2020). The anammox house: On the extracellular polymeric substances of anammox granular sludge.

Dueholm, M. K. D., Besteman, M., Zeuner, E. J., Riisgaard-Jensen, M., Nielsen, M. E., Vestergaard, S. Z., ... & Nielsen, P. H. (2023). Genetic potential for exopolysaccharide synthesis in activated sludge bacteria uncovered by genome-resolved metagenomics. Water Research, 229, 119485.

Felz, S., van Loosdrecht, M. C. M., & Lin, Y. (2019). Structural extracellular polymeric substances from aerobic granular sludge (Doctoral dissertation, Delft University of Technology).

Petriglieri, F., Singleton, C., Peces, M., Petersen, J. F., Nierychlo, M., & Nielsen, P. H. (2021). "Candidatus Dechloromonas phosphoritropha" and "Ca. D. phosphorivorans", novel polyphosphate accumulating organisms abundant in wastewater treatment systems. The ISME Journal, 1-10.

Seviour, T., Derlon, N., Dueholm, M. S., Flemming, H. C., Girbal-Neuhauser, E., Horn, H., ... & Lin, Y. (2019). Extracellular polymeric substances of biofilms: Suffering from an identity crisis. Water research, 151, 1-7.

Tomas Martinez, S. (2023). Extracellular Polymeric Substances of" Candidatus Accumulibacter": Composition, application and turnover.

Wang, J. Q., Yin, J. Y., Nie, S. P., & Xie, M. Y. (2021). A review of NMR analysis in polysaccharide structure and conformation: Progress, challenge and perspective. Food Research International, 143, 110290.

Curriculum Vitae

Stefan de Bruin was born on the 6th of November in 1992 in San José (Costa Rica), but for the majority of his life grew up in Wageningen, The Netherlands. In 2016 he received his bachelor degree of Life Science and Technology from Leiden University and TU Delft. After which he started and completed the MSc Life Science and Technology at the TU Delft in 2019. During his master program he got interested in the study of extracellular polymeric substances (EPS), ensuing him to do his MSc thesis project on how EPS influences the granule morphology, under the supervision of Yuemei Lin. This topic had piqued his interest so much that after an internship at Biothane Veolia Delft, working with anaerobic digesters, he decided to pursue the PhD that you see here.

Currently, he continues to research EPS, in the REThiNk project, which aims to valorize EPS from excess activated sludge created in wastewater treatment plants.

List of publications

De Bruin, S., Vasquez-Cardenas, D., Sarbu, S. M., Meysman, F. J. R., Sousa, D. Z., van Loosdrecht, M. C. M., & Lin, Y. (2022). Sulfated glycosaminoglycan-like polymers are present in an acidophilic biofilm from a sulfidic cave. *Science of the Total Environment*, *829*, 154472.

Chen, L. M., **de Bruin, S.**, Pronk, M., Sousa, D. Z., van Loosdrecht, M. C., & Lin, Y. (2023). Sialylation and Sulfation of anionic Glycoconjugates are common in the extracellular polymeric substances of both aerobic and anaerobic granular Sludges. *Environmental Science & Technology*, *57*(35), 13217-13225.

Doloman, A., **de Bruin, S.**, van Loosdrecht, M. C., Sousa, D. Z., & Lin, Y. (2024). Coupling extracellular glycan composition with metagenomic data in papermill and brewery anaerobic granular sludges. *Water Research*, *252*, 121240.

de Bruin, S., Hof, C., van Loosdrecht, M. C. M., Sousa, D. Z., Lin, Y. (2024). FT-IR micro-spectroscopy for imaging the extracellular matrix composition in biofilms bioRxiv 2024.08.26.609418; doi: https://doi.org/10.1101/2024.08.26.609418

Conference contributions

IWA biofilms 2022 conference, Phuket, Thailand (2022). Imaging the extracellular matrix in anaerobic granular sludge using FT-IR microscopy. *Oral*

Biofilms 10 conference, Leipzig, Germany (2022). Sulfated glycosaminoglycans are present in an acidophilic biofilm from a sulfidic cave. *Poster*

IWA Biofilms 2020 Virutal Conference (online), Notre Dame, USA (2020). Sulfated glycosaminoglycans are present in an acidophilic biofilm from a sulfidic cave. *Oral*

Acknowledgements

The PhD journey is (thankfully) not a solo journey, and there are always people along the way that have helped, guided, inspired or pushed me in certain ways. Here, I would like to thank all those people.

Firstly, I would like to thank my supervisors **Yuemei, Mark** and **Diana**. You have given me an interesting topic to dive into and

Yuemei, thank you for being understanding, patient and so approachable making it possible to come to talk whenever. It was always a pleasure to witness your enthusiasm for trying out stuff in the lab, which has also inspired me to try new things. During our meetings I have learned and enjoyed from the 'visual' way of thinking for stories and ideas.

Thank you, **Mark**, for being an unending source of ideas, point of views, and knowledge during the PhD. I would always leave our meeting with something unexpected. Furthermore, thank you for creating this diverse and open research group and emphasizing the importance of Friday drinks.

Diana, thank you for being my supervisor from a bit further, even though I did not come as often to Wageningen as I would have wanted in the end, I always enjoyed visiting the group at WUR and meeting with you. I especially enjoyed your relaxed approach during our meetings.

Additionally, I would like to thank everyone that helped or contributed to the thesis in one way or another. **Diana** thank you for writing the first paper together, and helping with the analysis and the story. It was quite a learning process for me, but I have learned a lot from your comments and writing style. **Lemin** it was very nice to work together on our paper. Especially the discussions we had during the writing have helped me a lot with understanding what we were actually doing. Of course, this would not have been possible without **Martin** and **Jitske's** analysis of the sialic acids. **Anna**, thank you very much for your patience and hard work on the metagenomics paper we wrote. It was not an easy task to do, as a lot of the analysis was new for us, or new in general, and thanks to your perseverance I think we made a nice story. The last chapter could not have been done without the help of **Astrid** with the cryotome slicing and **Marlies** with the set-up and training with the FT-IR micro-spectroscope.

None of this work, could have been done without the help of the technicians to help with analysis and keeping the lab running smoothly. Thank you: **Ben, Dita, Zita, Kevin, Patricia** and **Dirk**.

Furthermore I would like to thank all the students that agreed to do their thesis with me, within this project: **Kim**, **Manon**, **Jorn**, **Eecke** and **Carina**. I have learned a lot

from working together and we have found some interesting findings thanks to your hard work.

My paranymphs **Lemin** and **Timmy** of course deserve a special shout out. **Lemin**, I really loved and still love our coffee times, which I think started with science talk about EPS but is now filled with just chitchat. As a fellow E-P-S PhDer we naturally have been travelling together to a lot of conferences and there you have shown me 1) how super nice it is to travel in Thailand and 2) it is fine to take your time going to the airport. Thank you for both these insights. **Timmy**, I remember that a few years ago you asked me: "please can you make me your paranymph", I guess this is a classic be careful what you wish for kind of sentence, jaja. Thank you for sitting next to me, and putting up with my weird timings for asking you things *waiting until you put on your headphones* and being such a metiche in mine and other people's lives.

Other very important people during my PhD were: **Rodoula** and **Kiko**, my unofficial paranymphs. **Rodoula**, you really are a supergirl, you came into our office and inspired us, the boys and timmy, to dress better and behave better. Thank you for keeping our office in check while at the same time teaching us a little Greek like: "Ate Kiko", "Re Timmy". I also loved how we would tell each other our daily goals. **Kiko**, you have been such a source of inspiration for being so passionate about the things you do. I really liked us learning together about the countries of the world with the worldle game. Thank you for sharing so enthusiastically your successes, music, Catalan pride, and annoyances ("me cago en la ….").

I would like to thank these friends and fellow PhD's/EBTeans:

Ali, for providing us with fun facts during lunch and for being my roommate and diving/sauna/massage buddy during the EBT retreats and the Thailand trip, which were such chill times. **Nina**, for showing us how to dance the pepas song, and being Sam and me's personal pepas alarm during the Aalborg trip. **Sergio**, the S to our E-P, thank you for showing us all your moves when Toxic from Britney is on, and for taking us to the drag queen show, for sure that was an unforgettable moment for me. **David**, for your volume during lunches which made sure it was never quiet. Also for your wisdom about the thesis process, it always helped me to relativize my own PhD experience. **Maxim**, for being the calm chef during what might have been my most stressful moment during the PhD. **Sam**, for for being a fact checker of outrageous statements blurted out during lunch. Also thank you for not choosing this PhD topic, and passing it on to me! **Mariana**, your passion for Eurovision has also caused me to enjoy the song festival. A blessing and a curse… Thank you for always hosting these and other fabulous parties. **Jan,** thank you for joining me in doing weird challenges. **Angelos,** for your overall always cheerful demeanor and enthusiasm towards going out to get one more beer. **Venda**, thank

you for showing how to do these crazy beta breaks at DB. **Matteo,** thank you for thinking I am good at bouldering, you are so kind to give me an ego boost like that.

From the older generation **EBT**: **Hugo** for showing me around in Aalborg and your weird fermentations and cake creations, **Chris** for teaching me about Chinese knitting technique and gullibility, it was a good lesson, **Ingrid** for I think still my favourite PhD defence party and your chill vibes. **Philip,** thank you for your late night wisdom at the Oude Jan.

I would of course also like to thank the newer PhD's and members of **EBT**, you have made the lunches, Friday drinks, retreats and volleyball, football, bachata workshops and cake times super nice and welcome distractions: **Bea, Marit, Siem, Ramon, Jitske, Ji, Jelle, Linghan, Puck, Goncalo**, **Sebastian**, **Emine**, **Berdien, Yubo, Dinko, Samarpita** and of course **Natalya** for your enthusiasm with giving bachata, salsa and dancing workshops, which has taught me that I need to up my dancing game.

Outside of work I have been lucky to be a part of many varied friend groups which have pulled me out of the PhD bubble and into their world:

My **Nerd-Game-Gang** friends that started as a board sport/skateboard friend group but then morphed into a boardgame, bouldering, hiking, biking and whatnot friend group: **Alex**, **Kim, Ben, Bender, Cesc, Chris, Demi, Dion, Gijs, Gwen, Hidde, Karin, Lucia, Puck, Tim, Timo, Willemijn, Wouter, Yanni, Yoshi, Simon** and **Tristan**. Thank you for showing the nice dichotomy of partying/roadtripping/miniramping/skateboarding/faxemachining/wizardstaffing /hiking and boardgames/Magic/D&D/geeking out about technical gear and other nerdy stuff. I still love than when we are together we start to challenge each other to climb/jump into all kinds of things.

My friends and roommates from **JVB 35**, also known as **Huize boerenkool**: **Casper, Lisanne, Lucas, Lucia, Noor, Jaap, Tobias, Floor, Fréderique, Thomas, Emilie, Gerrit, Martine, Martijn K, Martijn S, Jesse, Louis, Els, Avelien, Stephan, Vera, Maurice, Floris, Deanne, Boris, Bas, Mette, Mathilde** and many more. Living together with you has taught me compassion for and understanding of others. Which is something that I keep on learning when we are together on our oud-huischgenoten weekends, or our vakantie bij de zon. Because of our holidays at the sun I have learned what the real bourgondisch leven is about, and I have really learned to appreciate cold rosé and kilo's of cheese. And even though we do not live together anymore, when we are together it still feels like home.

Lucia, for the majority of the PhD you were one of the most important people in my life. Thank you for showing your genuine empathic interest to other peoples and animals. I have learned a lot from that. You also emphasized the importance of a good work-life balance, listened to me ramble about work, was critical of my work

(in a good way). I am also grateful for the cosy and beautiful home we built together, with **Olli** (even though I am not sure I should put her here because she could be a terror).

My **Kubus** friends: **Bastian, Bas, Burger, Daan, Henk, Jelle, Joël, Lieuwe, Neuman, Rick, Rudo, Tiemen, Timo and Wijnand**, what a crazy long time have we been friends. Some of us have known each other for almost 20 years. It is nice and weird to see how we have changed from being little kids that eat waaay to much in between hours to being adults that talk about VVE's, children and marriage. Even though we still eat too much snacks when I sometimes come to huisavond on Thursday. You helped me get out of the Delft bubble and have taught me minimum required knowledge about a diversity of topics including travelling, living abroad, politics, economics and crazy outfit idea's during derde kerstdag.

No sería completo sin darles las gracias a mis padres, **Sytze** y **Shirley**, y a mi hermana **Natalia**. Ustedes siempre me han apoyado durante mi vida y durante el PhD. Siempre estaban interesados en el progreso del PhD. Y papa por suerte tambien me dio un poco de insight en la vida del professor. No estaría aquí sin ustedes.