THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Analysis of Heterogeneity in Polyglycerol Ester Surfactants

FRIDA BILÉN

Department of Chemistry and Chemical Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2026

Analysis of Heterogeneity in Polyglycerol Ester Surfactants
FRIDA BILÉN
ISBN 978-91-8103-472-1

Acknowledgements, dedications, and similar personal statements in this thesis, reflect the author's own views.

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5929
ISSN 0346-718X
https://doi.org/10.63959/chalmers.dt/5929

Department of Chemistry and Chemical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover: An illustration visualizing some examples of the molecules that can be present in heterogeneous polyglycerol ester surfactant products.

Printed by Chalmers Digitaltryck
Gothenburg, Sweden 2026

Analysis of Heterogeneity in Polyglycerol Ester Surfactants
FRIDA BILÉN
Department of Chemistry and Chemical Engineering
Chalmers University of Technology

Abstract

The properties and performance of materials are strongly influenced by their structure, yet many natural and synthetic materials consist of complex mixtures that vary in composition, size, or structure, collectively referred to as heterogeneity. Understanding this heterogeneity is essential for establishing reliable structure-property relationships, yet it is often difficult to characterize in detail, and its influence on material performance is therefore not always well understood. This thesis investigates heterogeneity from several perspectives, with a primary focus on polyglycerol ester surfactants, in which heterogeneity comes from both the polyglycerol headgroup and the fatty acid tail attachment.

A methodology based on quantitative and multidimensional NMR spectroscopy was developed to characterize and quantify the constitutional heterogeneity of commercial polyglycerol materials. The results showed that polyglycerols consist of mixtures of molecules containing structurally distinct glycerol subunits connected through different linkage patterns, and that their constitutional complexity increases with oligomer size. A simplified approach based on quantitative ^{1}H NMR was also demonstrated, offering a more rapid alternative for routine characterization.

The influence of heterogeneity on surfactant properties was investigated for polyglycerol ester surfactants, where a varying degree of esterification was considered. Heterogeneity was found to significantly affect solution behavior, adsorption kinetics, and foaming performance, while equilibrium surface tension was comparatively less sensitive to composition. In a complementary study, structurally well-defined branched triglycerol ester surfactants were used to establish structure-property relationships while minimizing molecular heterogeneity. This revealed unexpected effects of molecular architecture on self-assembly and interfacial behavior, highlighting the importance of headgroup structure and intermolecular interactions.

To explore data-driven approaches for characterization of heterogeneity more broadly, machine learning was applied to colloidal gold nanoparticles. Convolutional neural networks were trained to predict particle size distributions directly from UV–vis spectra, successfully extracting information about both average particle dimensions and polydispersity, and demonstrating the potential of machine learning as a tool for characterization of heterogeneous materials.

Overall, this thesis demonstrates that heterogeneity can be quantified using advanced experimental and data-driven approaches, and that heterogeneity can strongly influence surfactant performance.

Keywords: heterogeneity, polyglycerol, surfactants, polyglycerol esters, machine learning, NMR spectroscopy, gold nanoparticles.

List of publications

This thesis is based on the work contained in the following papers:

I. **Determination of polyglycerol substructures and their connectivity by multidimensional and quantitative NMR spectroscopy**
Frida Bilén[‡], Nasim Nahavandizadeh[‡], Moa Larsson, Sami Zeliouche, Romain Bordes, and Lars Evenäs.
Submitted

II. **Effect of molecular heterogeneity on properties for polyglycerol ester surfactants**
Frida Bilén, Moa Larsson, Lars Evenäs, Krister Holmberg, Romain Bordes.
Submitted

III. **Solution and interfacial behavior of branched polyglycerol ester surfactants**
Frida Bilén[‡], Eric Moilanen[‡], Moa Larsson, Krister Holmberg, Lars Evenäs, Romain Bordes.
Manuscript

IV. **Machine learning-based interpretation of optical properties of colloidal gold with convolutional neural networks**
Frida Bilén, Pernilla Ekborg-Tanner, Antoine Balzano, Michaël Ughetto, Robson Rosa da Silva, Hannes Schomaker, Paul Erhart, Kasper Moth-Poulsen, and Romain Bordes
Published in *The Journal of Physical Chemistry C*. **2024**, 128(33): 13909-13916.

[‡]The authors contributed equally to this work

Contribution report to the listed publications:

I. Shared main author. Part of the conceptualization. Co-responsible for data interpretation and visualization. Co-responsible for writing the first draft of the manuscript with the other main author.

II. Main author. Original idea and conceptualization. Responsible for the experimental work, except flash column chromatography and diffusion NMR. Responsible for writing the first draft of the manuscript.

III. Shared main author. Part of the conceptualization. Supervised the characterization part of the experimental work. Responsible for writing the first draft of the manuscript.

IV. Main author. Part of the conceptualization. Responsible for most of the data processing and machine learning work. Responsible for writing the first manuscript draft.

Abbreviations

1D	One-dimensional
2D	Two-dimensional
AI	Artificial Intelligence
AR	Aspect Ratio (AR=length/diameter)
AuNP	Gold Nanoparticle
CMC	Critical Micelle Concentration
CNN	Convolutional Neural Network
CPP	Critical Packing Parameter
DEPT	Distortionless Enhancement by Polarization Transfer
DLS	Dynamic Light Scattering
FTIR	Fourier transform infrared
HLB	Hydrophilic-Lipophilic Balance
HMBC	Heteronuclear Multiple Bond Correlation
HSQC	Heteronuclear Single Quantum Coherence
INADEQUATE	Incredible Natural Abundance DoublE QUAntum Transfer Experiment
ML	Machine Learning
MSE	Mean Squared Error
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
PEG	Polyethylene glycol
PG	Polyglycerol
QCM-D	Quartz Crystal Microbalance with Dissipation monitoring
ReLU	Rectified Linear Unit
SAM	Self-Assembled Monolayer
UV-vis	Ultraviolet-visible

Contents

1 Introduction

Few systems in nature or in the laboratory are truly uniform. Most materials, whether natural or synthetic, are composed of parts that differ from one another in some respect, a property known as heterogeneity. The term originates from the Greek *hetero* (different) and *genos* (kind), and stands in contrast to homogeneity, which describes systems in which the components are similar or uniformly distributed. In chemistry, the term is used in different situations, but often to describe systems containing multiple phases. For example, a heterogeneous reaction may involve reactants in different phases, such as solids, liquids, or gases, or occur at the surface of a solid catalyst [1]. A transparent one-phase solution, for example salt in water, is referred to as a homogeneous system, although it is composed of different ions and molecules, while a mixture of non-miscible oil and water gives a turbid appearance and would be referred to as heterogeneous.

In this thesis, however, the word heterogeneous is used to describe mixtures of molecules or particles that are similar, without being strictly identical. Many naturally occurring and synthetic materials fall into this category. Vegetable oils consist of mixtures of triglycerides containing fatty acids of different chain lengths and degrees of unsaturation. Polymers are often present as molecular weight distributions rather than single chain lengths. Likewise, colloidal nanoparticle systems commonly consist of distributions in particle sizes and shapes. Such heterogeneity can have a significant influence on material properties and performance.

One important class of heterogeneous materials is surfactants. Surfactants are surface-active molecules containing both a hydrophilic (water-loving) part and a hydrophobic or lipophilic (oil-loving) part. Their amphiphilic nature causes them to adsorb at interfaces and reduce interfacial tension between immiscible phases. This makes surfactants essential components in a wide range of products and applications, including detergents, shampoos, cosmetics, foods and in various industrial applications.

In many commercially available surfactants, neither the hydrophilic head group nor the hydrophobic tail group is molecularly pure. The tail group may consist of a distribution of alkyl chain lengths or degrees of unsaturation, while the head group may contain distributions in oligomer length or molecular structure. Furthermore, additional heterogeneity may be introduced during surfactant synthesis when the head and tail groups are combined, regarding tail positions and number of tails (degree of substitution). As a consequence, many surfactant products are not single molecular species but rather complex mixtures of molecules. Although heterogeneous

surfactant mixtures are widely used and often exhibit excellent performance, their complexity makes it challenging to establish quantitative structure-property relationships. It becomes difficult to determine which structural features are responsible for a given physicochemical property and how the various molecular species contribute to the overall performance of the material.

One surfactant class where these challenges are particularly pronounced is polyglycerol (PG) ester surfactants, which is the main focus of this thesis. PG esters are a type of nonionic surfactant of renewable origin, high biodegradability and low toxicity, and the interest for this group of surfactants has grown in recent years [2, 3]. PG esters consist of a hydrophobic fatty acid tail attached to a polyglycerol head group. The PG itself is a heterogeneous material produced by oligomerization of glycerol. During this process, glycerol units can connect in different ways, resulting in distributions of molecular sizes as well as constitutional structures [4]. Depending on how the glycerol units are connected, linear, branched, and cyclic structures can be formed. A few examples of PG structures are shown in Figure 1.

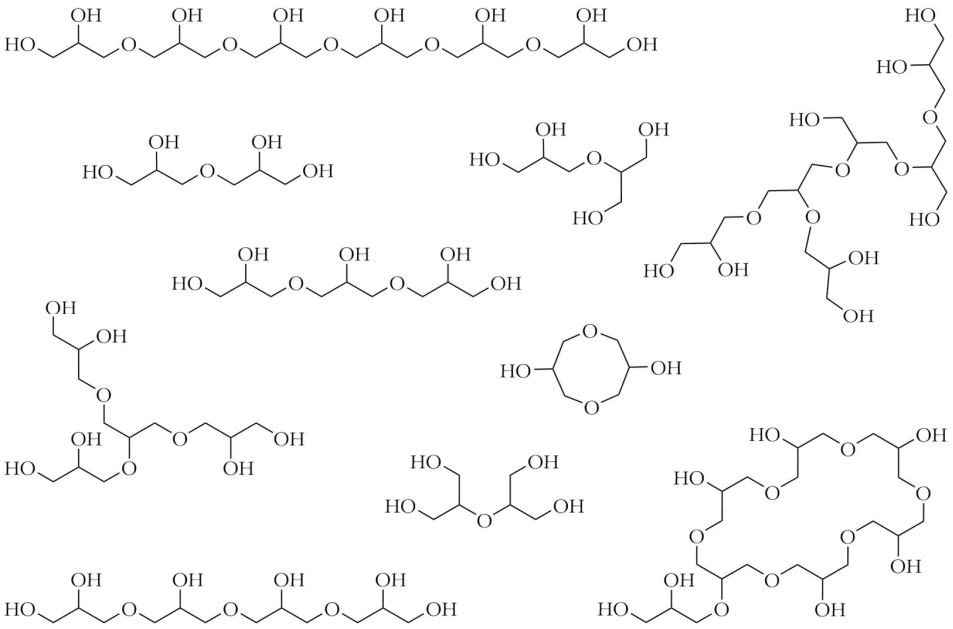

Figure 1. Examples of polyglycerol molecules, with different number of glycerol units, connected in different ways, forming linear, branched or cyclic structures.

The heterogeneity increases further during surfactant synthesis. Each PG molecule contains multiple hydroxyl groups that can act as reaction sites for fatty acid attachment. As a result, fatty acid chains can be attached at different positions on the PG molecule, and several fatty acid chains can be attached to the same PG, see Figure 2 for a few examples. Consequently, even a hypothetical monodisperse PG material would generate a heterogeneous surfactant mixture after esterification.

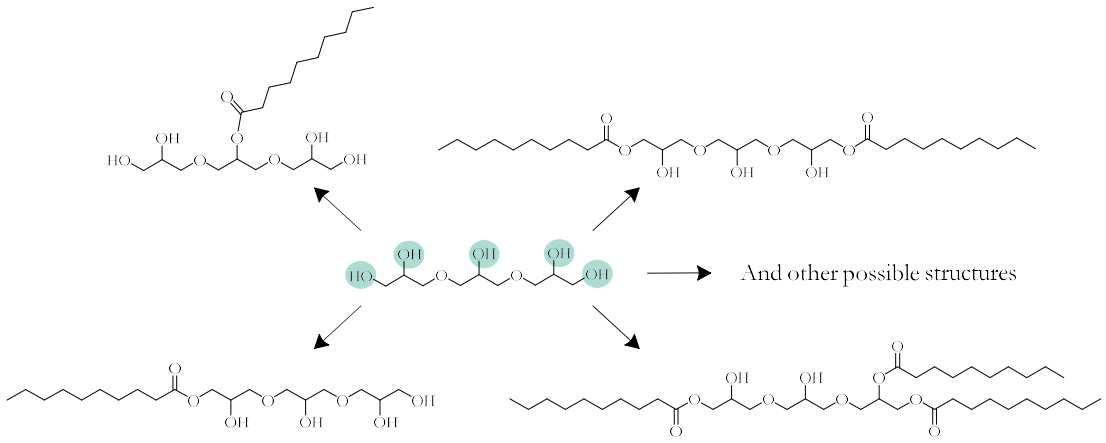

Figure 2. Schematic illustration of some possible polyglycerol ester structures. The multiple hydroxyl groups available for esterification allow fatty acid tails to attach at different positions and in varying numbers, generating many possible molecular structures.

Understanding the role of heterogeneity in such systems requires both characterization of the heterogeneity itself and investigation of how that heterogeneity influences material properties. Several complementary approaches can be used to address these challenges. One approach is to characterize and quantify the heterogeneity present in the raw materials and surfactant products. Another is to compare heterogeneous mixtures with purified fractions or to study structurally well-defined model compounds in order to isolate the effects of individual structural features.

Surfactants are, however, not the only important class of heterogeneous materials. There are many natural and manmade materials that are mixtures of molecules or particles, and colloidal gold is one of them. Gold nanoparticles often exist as distributions in size and shape, and these variations can influence the optical properties of the material [5]. Just as in surfactant systems, reliable characterization of the heterogeneity is essential for understanding and controlling performance.

The work presented in this thesis investigates heterogeneity from several different perspectives. The first part focuses on characterization of molecular heterogeneity in polyglycerol raw materials. The second part examines how heterogeneity influences the properties of polyglycerol ester surfactants and how structure-property relationships can be established despite the typical molecular complexity. The final part considers colloidal gold nanoparticles as a second heterogeneous material system and explores how machine learning can be used to extract information about size distributions from spectroscopic data. Together, these studies illustrate different strategies for characterizing and understanding heterogeneous materials.

1.1 Purpose and objectives

The purpose of this thesis was to deepen the understanding of heterogeneity in surface chemistry and to investigate methods for characterizing and analyzing heterogeneous systems. The primary focus was on polyglycerol-based surfactants, where heterogeneity was studied at the molecular scale, both at the raw-material level and in the final surfactant products. In addition, a data-driven approach for characterization of heterogeneity was investigated, applied to colloidal gold as a model system.

The thesis is based on four studies:

- In **Paper I**, the objective was to develop a methodology for characterization and quantification of the constitutional heterogeneity in polyglycerol materials using quantitative and multidimensional NMR spectroscopy. This work addresses the fundamental challenge of characterizing the raw material before meaningful structure–property relationships can be established.
- In **Paper II**, the objective was to investigate how heterogeneity in degree of esterification affects the properties of polyglycerol ester surfactants. Particular focus was placed on solution behavior, adsorption at the air-water interface, and foaming properties.
- In **Paper III**, the objective was to study structurally well-defined branched triglycerol ester surfactants in order to establish structure-property relationships while minimizing the effects of molecular heterogeneity. By controlling the structure, the contribution of specific molecular features to surfactant behavior could be investigated.
- In **Paper IV**, the objective was to investigate how machine learning can be used to characterize heterogeneity from ultraviolet-visible (UV–vis) spectra in colloidal gold nanoparticle systems. Convolutional neural network models were developed to predict particle size distributions directly from spectroscopic data, thereby exploring a data driven approach for analyzing heterogeneous materials.

2 Background

2.1 Heterogeneity in surfactants and their building blocks

2.1.1 About surfactants

Surfactants (surface active agents) are molecules with at least one hydrophilic (polar) part, called the head group, and one hydrophobic (non-polar) part, called the tail. The tail is often a hydrocarbon chain, branched or linear, typically containing 8-18 carbons [6], while the head group can vary considerably. Surfactants are divided into four categories according to the charge of the head group: nonionic, anionic, cationic, or zwitterionic. This thesis covers only nonionic surfactants, and the next section discusses several different types of nonionic surfactants in more detail.

The amphiphilic structure of surfactants makes them surface active, since the molecule is not fully soluble in either polar or nonpolar solvents and their hydrophilic and hydrophobic regions preferentially interact with different phases. As a result, surfactants tend to accumulate at interfaces, i.e. adsorb, where they can arrange themselves with their polar head towards the polar phase, and then non-polar tails towards the non-polar phase. As they adsorb at the interface, they lower the interfacial tension. Interfacial tension (or surface tension for the air-water interface specifically), arises from an asymmetry in attractive forces at the interface. In the bulk, a molecule experiences the same attractive forces in all directions, while at the surface, there is a net force towards the bulk. It can also be explained by that molecules at the surface have a higher free energy compared to molecules in the bulk, which drives an attempt to minimize the surface. As surfactants adsorb at the interface, the free energy, and therefore the surface tension, is lowered [7]. A schematic illustration of a few examples of surfactants at interfaces can be seen in Figure 3.

The surface active properties of surfactants are exploited in a wide range of applications, for example in cleaning products, where surfactants enable wetting of soiled surfaces and dispersal of dirt so it can be rinsed off with water. Other examples of applications include the use as

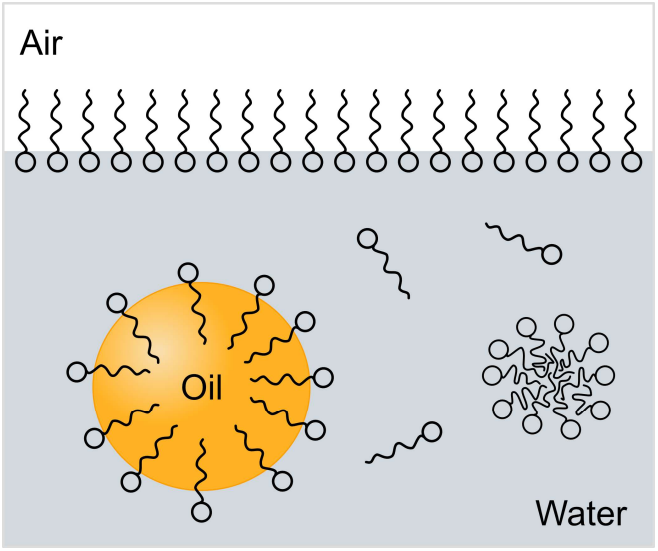

Figure 3. Schematic representation of surfactants, as unimers in solution, self-assembled into a spherical micelle, at the air-water interface, and at the oil-water interface as stabilizers for an emulsion droplet.

emulsifying agents in various industries and household products where an oil phase and a water phase need to mix, and as foam stabilizers, where surfactants accumulate at the air-water interface. The hydrophilic-lipophilic balance (HLB) of a surfactant, which is a measure of the balance between the polar and non-polar parts of the molecule, can give an indication of what the surfactant is suitable for in terms of applications or emulsion types [8].

Surfactants will also self-assemble in water to minimize their free energy, which means that the molecules will form aggregates with the insoluble part of the molecule inward, to minimize contact with the solvent. A common type of self-assembly is the spherical micelle. Figure 3 also shows a cross-sectional illustration of a spherical micelle, in which several surfactant molecules come together with their tail groups oriented toward the center and their polar head groups turned outward toward the water. It is not until the surfactant concentration reaches a certain point that the surfactants begin to self-assemble into aggregates, often micelles. This concentration is called the critical micelle concentration (CMC). Below the CMC, surfactants are present as free unimers in solution or accumulate at interfaces. Once the CMC is reached, however, any additional surfactant added to the solution forms micelles rather than contributing further to the unimer concentration. The CMC is an important characteristic of a surfactant, reflecting its tendency to self-assemble and marking the concentration above which micelles begin to form. It can be determined, for example, by measuring the surface tension of an aqueous solution of surfactant at different concentrations, and plotting the surface tension against the logarithm of the concentration, see Figure 4. An inflection point is seen in the curve at the CMC, and above the CMC the surface tension remains close to unaffected by the surfactant concentration, which is referred to as the surface tension at the plateau. A low surface tension at the plateau indicates a high surface activity of the surfactant [9]. Although the surface tension is expected to become approximately concentration-independent above the CMC, in practice a slight drift may remain because of slow interfacial equilibration, surfactant polydispersity, impurities, or concentration-dependent changes in the composition and structure of the aggregates and adsorbed layer.

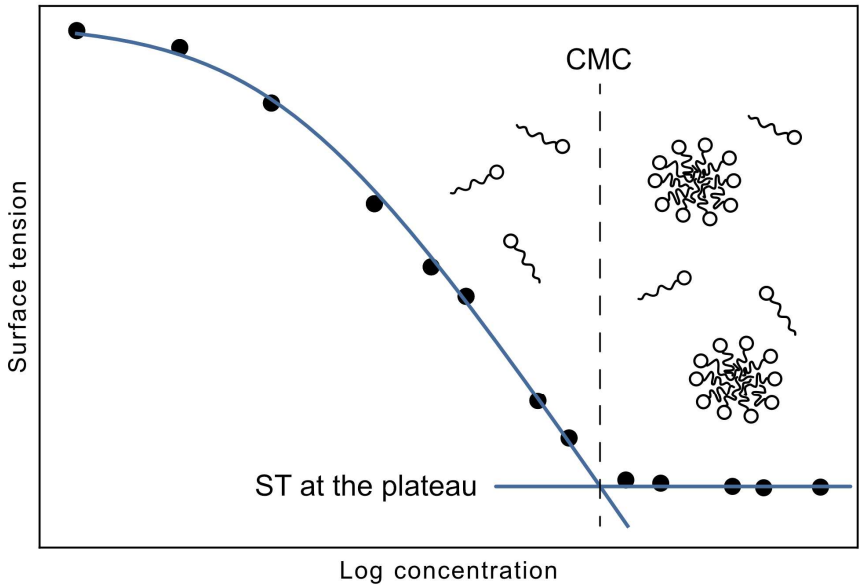

Figure 4. An example surface tension isotherm, where the CMC is marked as the inflection point of the curve, above which the surface tension is close to unaffected by the surfactant concentration, and this surface tension (ST) is called the ST at the plateau.

The CMC depends largely on the length of the hydrophobic tail. The relationship between CMC and tail length can be described by:

$$\log CMC = A - Bn \qquad (1)$$

where A and B are constants, and n is the number of carbons in the tail. The slope, B, is typically 0.5 for nonionic surfactants, and 0.3 for ionic surfactants [10, 11].

The spherical micelle described above is just one type of surfactant self-assembly. Several other self-assembled structures exist, and the preferred structure depends on the geometry of the surfactant molecule, which can be estimated using the critical packing parameter (CPP):

$$CPP = \frac{v}{a \cdot l_{max}} \qquad (2)$$

where v is the volume of the surfactant tail, a is the effective head group area, and l_{max} is the length of the surfactant tail [12, 13]. A small CPP (<1/3) typically favors spherical micelles, due to the comparatively large head group area relative to the tail volume. At slightly larger CPP (1/3-1/2), rod-like self-assemblies are formed, and as the CPP approaches 1, the surfactant molecule adopts a more cylindrical shape, which promotes larger self-assembled structures of low curvature, such as lamellar phases or vesicles. A CPP greater than 1 favors reverse self-assemblies, in which the head groups point toward the center of the aggregate and the tails extend outward toward a non-polar solvent [13].

It should be noted that the effective head group area is not just the geometrical size of the head group, but also depends on the interactions with other surfactants and the solvent. For example, ethoxylated surfactants typically undergo a conformational change upon heating, in which the polyoxyethylene head group becomes less hydrated at higher temperatures, giving a smaller effective head group area. This promotes micellar growth, causing the solution to turn cloudy as

the larger self-assemblies form and begin to scatter light. The temperature at which this occurs for a 1 wt% solution is called the cloud point [14].

2.1.2 Different building blocks for surfactants synthesis and their heterogeneity

As discussed previously, surfactants consist of a hydrophobic and a hydrophilic part. Both of these components can originate from a range of raw materials and, in many cases, neither the head group nor the tail group exists as a single well-defined molecular species. Instead, surfactants are frequently produced from raw materials that are inherently heterogeneous, leading to distributions in chain length, molecular size, branching, degree of substitution, and other structural features of the molecules.

2.1.2.1 Heterogeneity in hydrophobic raw materials

A major source of hydrophobic raw materials for surfactants is naturally occurring fats and oils. These materials consist primarily of triglycerides, where the fatty acids vary in chain length and degree of unsaturation. Different natural sources have different fatty acid compositions, for example coconut oil with mainly medium chain fatty acids (C12-C14) and tallow fat with mostly long-chain fatty acids (C16-C18, including a large fraction of oleic acid with an unsaturation) [15]. Therefore, by choosing the source of the raw material one can influence the composition to some extent, but essentially all fats and oils of natural origin are mixtures of different molecular species [15]. In addition, vegetable oil composition from a certain species can be influenced by environmental factors, such as temperature [16] and drought stress [17] during plant growth.

The triglycerides can be used to make different tail groups for surfactant synthesis, such as fatty acids and fatty alcohols. The heterogeneity in the hydrophobic chains of the surfactant then remains through to these hydrophobic building blocks. There are, however, some methods to purify the tail groups, such as fractional distillation, allowing surfactants to be prepared from more narrowly defined hydrophobic raw materials [15].

In some cases, the fats and oils themselves can be used directly in surfactant synthesis. This introduces additional complexity because the final product may contain also unreacted triglycerides, diglycerides, monoglycerides, free fatty acids, and glycerol. These components can either participate in the synthesis or remain as byproducts, leading to products that are substantially more complex than those produced from purified fatty acids or fatty alcohols [18].

2.1.2.2 Heterogeneity in hydrophilic building blocks

The hydrophilic part of nonionic surfactants can also introduce substantial heterogeneity. Different classes of nonionic surfactants differ not only in the chemical nature of their head groups but also in the extent and origin of their molecular distributions. Some examples of nonionic head groups with considerable heterogeneity are presented below.

Ethoxylated surfactants

Ethoxylated surfactants are the most widely used class of nonionic surfactants. Their hydrophilic component consists of polyoxyethylene chains, which can be combined with various hydrophobic groups such as fatty alcohols or fatty acids to make surfactants. The ethoxylated head groups are always a distribution of homologue lengths rather than a single size. The exact distribution depends on several factors, including the identity of the hydrophobic group, the amount of ethylene oxide used during synthesis, and the catalyst [19]. Although the average

number of ethylene oxide units can be controlled accurately, the product remains a distribution of oligomer lengths [6].

The width of the distribution can be influenced by the type of catalyst, and certain catalysts produce narrower distributions than others [6]. Such narrow-distribution or "peaked" ethoxylates can provide advantages such as improved performance, increased water solubility, lower viscosity, more efficient wetting and less skin irritation compared with broader distributions [6, 11, 19]. The distribution can also be narrowed through separation techniques, such as distillation, although the products still remain distributions rather than discrete molecular species [19].

In polyoxyethylene esters, an additional level of heterogeneity arises. They are prepared by reacting fatty acids with polyethylene glycol (PEG), allowing the attachment of the tails on two sites. Once one hydroxyl group from the PEG has been esterified, the second hydroxyl group at the other extremity of the molecule can react with an additional fatty acid molecule, giving rise to diesters. Consequently, the final product may contain a mixture of monoesters, diesters, and unreacted PEG chains, in addition to the heterogeneity already introduced by the chain-length distribution in the ethoxylated head group and any heterogeneity in the fatty acid feedstock. Monoester formation therefore often requires use of a large excess of PEG, followed by a removal of excess material after synthesis [18].

Alkyl polyglycosides

Alkyl polyglycosides contain hydrophilic head groups composed of varying numbers of glycoside units. As a result, the head group exists as a size distribution. In addition, the hydrophobic part is commonly derived from fatty alcohol mixtures, introducing another level of heterogeneity as mentioned above [20].

Sorbitan esters and polysorbates

Sorbitan ester surfactants are produced from sorbitol-derived materials that themselves consist of complex mixtures containing sorbitol, various sorbitan isomers, and isosorbide. Subsequent esterification introduces further complexity through variations in the number and position of fatty acid chains, resulting in highly heterogeneous products [15].

To improve water solubility, sorbitan esters are often ethoxylated, and these ethoxylated sorbitan esters are called polysorbates [15]. In addition to the heterogeneity already present in sorbitan esters, then additional heterogeneity is introduced to polysorbates through the distribution of ethoxylate chain lengths and positions [21-23].

Glycerol esters

Glycerol esters may be prepared either by reaction of glycerol with fats or oils, or by direct esterification of glycerol and fatty acids. Both routes produce mixtures containing species with different numbers of fatty acid chains and different substitution positions on the glycerol. Similarly to the other surfactant types, if mixed fatty acid feedstocks are used, additional heterogeneity arises from variations in chain length and unsaturation [15].

Polyglycerol esters

Polyglycerol (PG) esters represent one of the most structurally complex classes of nonionic surfactants and are therefore of particular relevance to this thesis. They are similar in structure to the above-mentioned glycerol esters, but the use of glycerol as a surfactant head group is sometimes insufficient to achieve and control the desired hydrophilic-lipophilic balance [4]. By oligomerizing glycerol into polyglycerol, the hydrophilicity can be increased and tuned more precisely. Polyglycerol used in surfactants is most commonly produced through oligomerization of glycerol itself, although alternative monomers can also be employed [4].

Glycerol contains three hydroxyl groups, and although the primary hydroxyl groups are more reactive than the secondary hydroxyl group [24, 25], all three can participate in the oligomerization reaction. As a result, glycerol units can connect in multiple ways, producing linear, branched, and even cyclic structures [4], of which a few examples were shown previously in Figure 1. Furthermore, the oligomerization generates a distribution of molecular sizes with different numbers of glycerol units. Thus, similar to ethoxylated surfactants, polyglycerol-based surfactants contain a distribution of head-group sizes, but are more constitutionally complex in comparison, as the ethoxylated oligomers are linear. The degree of polymerization and degree of branching of PG can be influenced through catalyst selection and reaction conditions [24, 26, 27], but the resulting material remains a complex mixture of molecules.

The subsequent esterification of PG introduces additional levels of heterogeneity. Because a PG molecule contains multiple hydroxyl groups, the fatty acid chains can attach at various positions. Furthermore, several fatty acid chains can attach to the same molecule, producing a mixture of monoesters, diesters, triesters, and higher esters. Compared to ethoxylated and glycerol esters, where only a limited number of attachment sites are available, polyglycerol esters offer a substantially larger number of possible substitution possibilities. The number of possible structures increases with increasing degree of polymerization.

Finally, if the fatty acid feedstock itself contains a distribution of chain lengths or degrees of unsaturation, as was discussed above, an additional source of heterogeneity is introduced. Therefore, while some of the heterogeneities presented here can be partially tuned via synthesis conditions, polyglycerol ester surfactants can exhibit heterogeneity arising from the hydrophobic raw material, the hydrophilic raw material, and the esterification process itself, leading to a highly complex mixture of surface active species.

2.1.2.3 Heterogeneity in surfactants: Challenges and opportunities

The examples above illustrate that heterogeneity can originate from the hydrophobic part, the hydrophilic part, or both. Additional heterogeneity may also be introduced during surfactant synthesis through side reactions, variations in degree of substitution, or differences in tail attachment positions.

Importantly, heterogeneity is not necessarily a problem. Commercial surfactants are used successfully as heterogeneous mixtures. In some cases, heterogeneity may even be beneficial. One study on alcohol ethoxylate surfactants compared a pure homologue to two technical surfactant products with broad distributions of the head group, and showed that the technical products could reduce surface tension more [28]. One study that compared PG esters with broad

and narrow distribution of the PG head group showed that the product with the widest distribution could lower the surface tension more, and had better emulsification properties, which was attributed to a denser packing of the surfactants at the interface [29].

Nevertheless, heterogeneity also presents several challenges. First, it may reduce atom economy by generating species that do not contribute to the desired performance while increasing the complexity of the product. Second, replacing suppliers, or even batches, of a raw material or surfactant product can alter product performance, because differences may exist in the underlying molecular distributions [23, 30]. This highlights the importance of detailed characterization of the raw materials and of the surfactants.

Heterogeneity also complicates the establishment of structure-property relationships. Understanding how and why a surfactant behaves as it does become considerably more difficult when the material consists of large distributions of molecular species. Several strategies can be used to address this challenge. One approach is to characterize and quantify the heterogeneity present in the material, as explored in **Paper I**. Another is to fractionate or purify complex products and compare the properties of the purified fractions with those of the heterogeneous mixtures, as explored in **Paper II**. A third strategy is to synthesize and study well-defined model surfactants that allow the role of individual structural features to be isolated, as demonstrated in **Paper III**. These approaches are discussed throughout this thesis.

2.2 Heterogeneity in colloidal gold

Another heterogeneous material that will be discussed in this thesis is colloidal gold, or gold nanoparticles (AuNPs), which are used in a range of applications. As is discussed further below, AuNPs absorb light of certain wavelengths, and this property is exploited in photothermal cancer therapy and antibacterial materials, where the absorbed energy is converted into heat that can kill cancer cells or bacteria [31-35]. The absorbance of light can also be used in biosensing, which relies on the fact that the wavelength of the absorbed or scattered light depends on the local environment around the nanoparticles, and how closely the nanoparticles are spaced relative to one another [31, 32, 36, 37].

Gold nanoparticles are so small that the properties of the material begin to deviate from those of bulk gold. At the nanoscale, the very high surface-to-volume ratio has a marked effect on material properties [38]. One phenomenon that emerges at this scale is localized surface plasmon resonance (LSPR). Upon irradiation with light of a certain resonance wavelength, the light is absorbed and causes oscillation of the electrons at the surface of the gold nanoparticle, see Figure 5. This absorption gives rise to a characteristic band in the UV-vis spectrum, and the wavelength at which resonance occurs depends on the size and shape of the nanoparticles. Spherical particles, being symmetric in all dimensions, typically show a single absorption band, whereas rod-shaped particles exhibit both a longitudinal and a transverse oscillation, along and perpendicular to the rod's long axis, respectively, giving rise to two distinct peaks in the spectrum [38]. There are nanoparticles with a number of different shapes and materials that have plasmonic properties, and the environment surrounding the particles also affects the optical properties of the particles [32]. This thesis is however limited to spherical and rod-shaped nanoparticles of gold in aqueous environments.

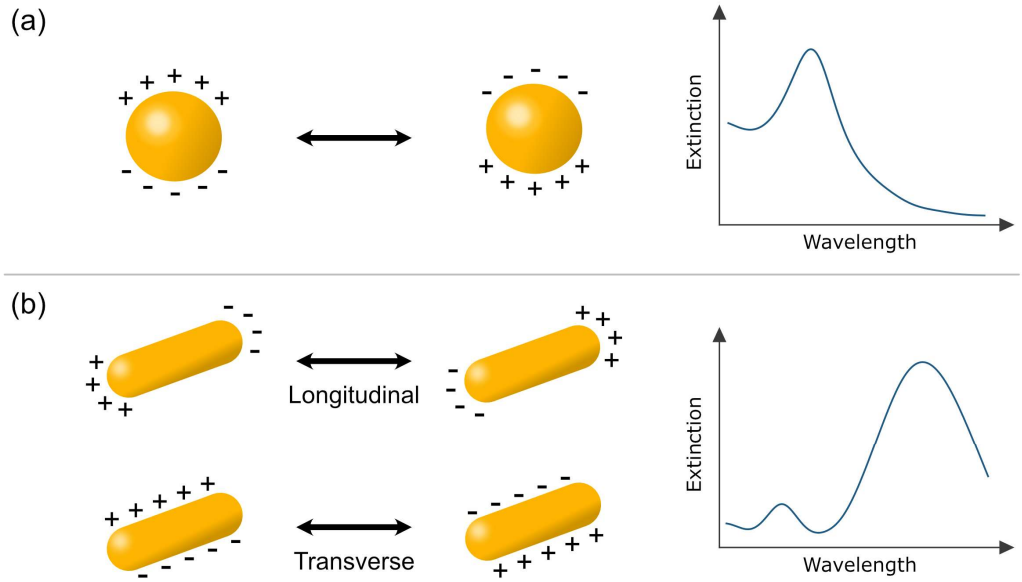

Figure 5. Illustration of the localized surface plasmon resonance in (a) spheres and (b) rods.

Beyond material and shape, particle size also affects LSPR [36, 38, 39]. For spherical nanoparticles, the wavelength of absorbed light is related to particle diameter [36, 38], while for gold nanorods it is proportional to the aspect ratio (AR, the ratio of length to diameter) [31]. Particle size also affects the width of the absorption band for different sizes of AuNPs [40]. As nanoparticle size increases, the contribution of scattering also increases [39], adding a further size-dependent effect on the resulting extinction (extinction being the sum of absorption and scattering). The size distribution or heterogeneity of gold nanoparticles therefore has a direct effect on their optical properties. For applications that depend on these properties, such as sensing and photothermal therapy, heterogeneity could consequently also affect performance. One example of this is that the polydispersity in particle size distribution has been shown to affect performance when AuNPs are used as sensors for solvent refractive index [37]. The size and aspect ratio have also been shown to affect the cellular uptake of AuNPs [41], which is an important consideration for biomedical applications. Taken together, this means that the size and shape of AuNP are central features that require careful evaluation.

Gold nanoparticles are typically synthesized by reduction of a gold salt in solution, where stabilizing ligands are used to prevent the formed gold nanoparticles from aggregating and to control the particle growth [32]. As mentioned above, it is important to control the size of the synthesized nanoparticles, and the particle size can be adjusted for example by varying the ratios between gold ions, reducing agents and stabilizing ligands [32]. There are methods, for example lithographic techniques, that can produce AuNPs with well-defined sizes and shape, but these techniques are experimentally demanding [31]. There are continuous advancements in conventional nanoparticle synthesis that improves the control of shape and size, but there are complex mechanisms behind the nanoparticle growth which makes precise control over the shape and size challenging [42]. The synthesis of gold nanoparticles can therefore yield a heterogeneous mixture of sizes with varying degrees of polydispersity, and particles can also aggregate over time, introducing additional heterogeneity in the sample [5]. It is therefore

important to be able to characterize this heterogeneity, for example to better fine-tune reaction conditions and achieve the desired properties.

The size distribution and shape of AuNPs are commonly characterized using electron microscopy [36]. However, this approach is typically time-consuming, requires expensive and complex instrumentation, and only samples a small fraction of the particles present. An alternative technique is UV-vis spectroscopy, since, as noted above, light absorption is directly related to particle size and shape. Several groups have developed equations or models that can be used to correlate the absorbed wavelength or width of the absorption band with the particle diameter or aspect ratio [43-47]. It is, however, difficult to draw conclusions about heterogeneity from UV-vis spectra alone, as this method typically yields only an average particle size. The spectra do contain information about the underlying size distribution, but they can be difficult to interpret. One approach to extract this information can be through machine learning assisted interpretation of the spectra, which is discussed in the next section of this thesis.

2.3 Machine learning as a tool to evaluate heterogeneity

Machine learning (ML), a subcategory of artificial intelligence, is a set of algorithms that can learn from and make predictions or decisions based on data [48]. This means that a ML model can learn directly from the data, without any preexisting knowledge about underlying physical or chemical phenomena [49]. ML is especially useful in situations where there is a lot of data to interpret, which would be time consuming for a human to assess and draw conclusions from [48], for automation of analysis processes [50], or in situations where there are small nuances in the data that are otherwise difficult to catch. This third trait of ML makes it promising for the study of heterogeneous materials.

There are three main types of machine learning algorithms: supervised, unsupervised and reinforcement learning. This thesis will cover only supervised learning, in which the model learns correlations based on data pairs: input and output data. Simply put, the model studies a large number of input-output pairs, and learns what the output should be based on the input. Then, when presented with previously unseen input, it makes a prediction of what the output should be. To be able to learn the correlation between the input and output, a large amount of data is needed, which is a considerable challenge in some applications [48].

2.3.1 Deep learning and neural networks

Deep learning is a subfield of ML in which the model is built as an artificial neural network with several layers: a structure loosely inspired by how neurons in the brain are connected. A neural network consists of layers of simple computational units ("neurons"), each of which takes in numbers, combines them mathematically, and passes a new number onward to the next layer. By stacking many such layers, the network can build up increasingly complex representations of patterns in the data. This lets neural networks learn subtle correlations directly from data, but requires large amounts of training data, since the network has many internal parameters ("weights") that must be tuned from examples [51, 52].

In this thesis, a type of model called Convolutional Neural Network (CNN) has been used, which is most known for its outstanding performance in image recognition. CNNs, and other types of neural network models, are especially good for complex, array-like data such as images (two-dimensional (2D) data) or spectra (one-dimensional (1D) data), as they can typically be used

directly without much need for data preprocessing or dimensionality reduction [52, 53]. Compared to fully connected neural networks, the CNN model focuses on neighboring datapoints, which optimizes it for finding patterns in images, such as vertical or horizontal lines, and in extension, various more complex patterns [51, 52].

2.3.2 The building blocks of a CNN

A neural network is organized into an input layer (where the raw data enters), one or more hidden layers (where the data is progressively transformed), and an output layer (which gives the final prediction) [51].

The hidden layers in a CNN typically include [51, 52]:

- Convolutional layers, which scan across the input data using a small sliding window called a kernel (or filter). At each position, the kernel combines its values with the underlying data to produce one output value that highlights a local pattern. The filter size determines how many neighboring datapoints the kernel considers at once, and a layer typically uses many filters in parallel to detect different patterns.
- Max-pooling layers, which typically come after convolutional layers and simplify the data. They keep only the largest value within small windows of the data (defined by a pool size), moving across the data in steps defined by the stride. Pooling layers help reduce the number of parameters, which speeds up the computations, and makes the model less sensitive to small variations in the input.
- Fully connected (dense) layers, in which every neuron is connected to every neuron in the previous layer. These are typically used toward the end of the network to combine the patterns detected earlier into a final prediction.

After each layer's calculations, an activation function is applied to each neuron's output. This introduces non-linearity into the network, which is essential for allowing it to learn complex, non-linear relationships rather than being limited to simple linear combinations of the input. A commonly used activation function, also used in this thesis, is the rectified linear unit (ReLU), which sets any negative value to zero and leaves positive values unchanged. Other activation functions exist and their selection is typically part of the optimization process of the network architecture [51].

2.3.3 Training a neural network

Before a network can make useful predictions, it must be trained. By showing it many input-output examples it can adjust, in an iterative manner, its internal weights to minimize the difference between its predictions and the known, correct outputs. This difference is quantified by a loss function (for example the mean squared error (MSE)), which gives a number describing how far off the model's predictions are [51].

Training proceeds using an optimization algorithm, which decides how to adjust the network's weights in order to reduce the loss (in this thesis, we used a commonly used algorithm called Adam [54]). Rather than processing the entire training dataset at once, the data is split into smaller groups called batches, and the network updates its weights after processing each batch. One full pass through all the batches, i.e. through the entire training dataset, is called an epoch. A

network is typically trained over many epochs, repeatedly cycling through the training data, so that it can gradually refine its weights [51].

To evaluate how well training is progressing and to avoid the model simply memorizing the training examples (a problem called overfitting), the available data is typically split into three parts: a training set, used to update the weights, a validation set, used during training to monitor performance on data not directly used for weight updates, and a test set, kept completely separate and used only after training is complete, to give an unbiased evaluation of the final model's performance on entirely unseen data [49].

2.3.4 Machine learning for characterization of gold nanoparticles

The previous sections described ML, and the basics of how a supervised CNN model is trained and used. But can it be used to study heterogeneity of materials? Heterogeneity has been characterized with ML previously, for example to find the molecular weight distribution of polymers from their rheological properties [55]. ML has also been used to determine the compositions of chemical mixtures from their Fourier transform infrared (FTIR) spectra [56]. In **Paper IV**, we explored the use of ML for characterization of the heterogeneity of gold nanoparticles. As was mentioned in section 2.2, the characterization of the size and shape of AuNPs is very important, as these parameters are directly related to their optical properties, and ML is one possible method to use for this characterization. ML enables an automated analysis which can be useful in for example flow synthesis of nanoparticles, where an inline characterization of the sizes could be used for real-time fine tuning of reaction parameters. As there is a correlation between the particle size and their optical response for the plasmonic AuNPs, ML can be used to extraction information about the nanoparticle sizes from UV-vis data. This has been done previously for prediction of average particle sizes either directly from spectra, or from parameters extracted from spectra, such as the peak position or width [57-59]. There are also examples of various ML algorithms used for prediction of the size distribution, i.e. a measure of the heterogeneity, for gold nanoparticles [60, 61], although these used various descriptors of the UV-vis spectra as input for their models, rather than the full spectra. In **Paper IV**, we show that by using a 1D CNN model, the full spectra can be used as input, and that the ML model then can extract detailed information about the polydispersity of the particle mixtures, for spherical and rod-shaped AuNPs. The CNN model proves to be powerful for identification of patterns relevant for these predictions.

3 Methodologies

In this section, the materials and methodologies that are the most relevant for this thesis are described.

3.1 Materials

The materials evaluated in this thesis are four polyglycerol materials and six polyglycerol ester surfactant samples.

The polyglycerol samples (PG2, PG3, PG4, and PG6) were obtained from Spiga Nord S.p.A. and have increasing degree of polymerization and polydispersity. The number in each sample name indicates the approximate size of the molecule in terms of the number of glycerol units, but each material is in fact a mixture with a distribution of sizes.

The six polyglycerol ester surfactants can be divided into two groups according to their synthesis paths, as described in Figure 6. Group (a) were made in a one-step synthesis from heterogeneous mixtures of PG, with non-specific tail attachments, while group (b) were made through a multi-step procedure to yield surfactants of well-defined molecular structures.

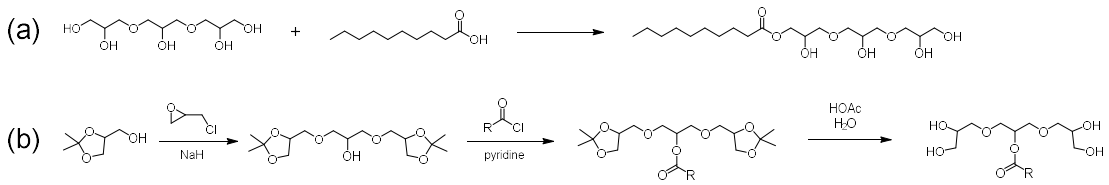

Figure 6. Synthesis paths for the two groups of polyglycerol ester surfactants studied in this thesis: (a) Heterogeneous mixtures synthesized through a one-step, solvent-free synthesis between decanoic acid and the heterogeneous PG3, resulting in a mix of head group structures, sizes and varying degree of esterification. (b) Well-defined triglycerol esters synthesized through a multi-step process, to selectively attach the tail (with 10, 12 or 14 carbons) to the central hydroxyl of the triglycerol.

The surfactants in the two groups can be described as follows:

(a) Heterogeneous mixtures of esters synthesized from PG3 and decanoic acid in a solvent-free, one-step reaction, as described in **Paper II**. The synthesis path did not provide control over tail attachment, and the tails could attach at different positions on the PG molecule, or with multiple tails attached to the same head group. To promote monoester formation, PG was used in excess, at either a 2:1 or 5:1 PG to fatty acid ratio, where the 5:1 ratio was expected to yield a higher monoester content. Two of the surfactant samples were used as obtained, after extracting the

esters from the excess PG with dichloromethane, and are referred to as 2:1 and 5:1 according to their respective PG:fatty acid reaction ratio. A third sample, referred to as ME (for monoester), was synthesized under the same conditions as the 5:1 sample but was subsequently purified by flash column chromatography to isolate the monoester fraction (details of the separation can be found in **Paper II**).

(b) Branched triglycerol esters of well-defined structure, with 10, 12, or 14 carbons in the hydrophobic tail (samples G3C10, G3C12, and G3C14, respectively). Detailed descriptions of these materials and their syntheses are provided in **Paper III**.

3.2 NMR spectroscopy

Some atomic nuclei, such as ^{1}H and ^{13}C, possess a nonzero nuclear spin, which gives rise to a nuclear magnetic moment. This makes it possible to study them with nuclear magnetic resonance (NMR) spectroscopy. When placed in a strong magnetic field, spin-½ nuclei (e.g. ^{1}H and ^{13}C) can occupy two discrete energy levels. The nuclear magnetic moments precess around the direction of the magnetic field at a characteristic frequency, called the Larmor frequency. When an applied radiofrequency pulse contains a frequency matching the Larmor frequency, resonance occurs and the nuclear spin system is perturbed from its equilibrium state. After the pulse, the nuclear magnetization relaxes back towards equilibrium, while the NMR instrument detects a signal containing information about the resonance frequencies. This signal is then used to construct the NMR spectrum. The frequency at which resonance occurs is affected by the chemical environment of the nucleus, resulting in signals at different frequencies for nuclei in different chemical environments. By looking at the chemical shift (the peak position) one can therefore obtain information about these environments, which is useful when determining the structure of a molecule [62].

When observing quantitative spectra, the integrals of the peaks can also be used to get information about the relative amounts of each nucleus type, as the peak area is proportional to the number of nuclei giving rise to the signal. For both ^{1}H and ^{13}C NMR, appropriate experimental conditions are required for quantitative spectra, although quantitative ^{13}C NMR requires particular consideration. To obtain a quantitative spectrum, the delay between successive pulses must be long enough, relative to the relaxation time, to allow sufficient relaxation towards the equilibrium state before the next pulse. Due to the low natural abundance of ^{13}C (ca 1%), ^{13}C NMR has relatively low sensitivity, and spectra are normally recorded with ^{1}H broadband decoupling to remove ^{13}C-^{1}H couplings and increase sensitivity. Continuous ^{1}H decoupling also gives rise to so-called nuclear Overhauser effect (NOE) enhancement of the ^{13}C signals. However, the magnitude of this enhancement varies between different carbon environments and the resulting signal intensities are therefore not quantitative. To obtain a quantitative ^{13}C NMR spectrum, inverse-gated decoupling can therefore be used, where decoupling is applied during acquisition but not during the relaxation delay, suppressing the NOE while retaining ^{1}H-decoupled ^{13}C signals [62].

To determine the molecular structures from the ^{1}H or ^{13}C spectra, one first has to assign the different peaks to what type of structure they arise from. As already mentioned, the chemical shift gives very useful information as it is affected by the chemical environment around the nucleus. In ^{1}H NMR, the splitting of a peak is also informative, as it arises from spin-spin coupling (*J*-coupling) to other nuclei through chemical bonds and provides information about their number and coupling relationships. For ^{13}C-NMR, the experiment is typically ^{1}H decoupled,

removing ^{13}C-^{1}H *J*-couplings and simplifying the spectrum, and the peaks therefore show up as singlets, without *J*-coupling to the surrounding hydrogens. Information about the number of hydrogens directly attached to a carbon can instead be determined from DEPT (Distortionless Enhancement by Polarization Transfer) experiments. Different DEPT experiments, commonly DEPT-45, DEPT-90 and DEPT-135, can be used to distinguish between CH, CH_2 and CH_3 carbons. One example is DEPT-135, which shows CH and CH_3 as positive signals, and CH_2 as negative signals [62].

Two-dimensional (2D) NMR techniques are also useful for peak assignment. 2D NMR experiments plot signal intensity against two frequency axes, with cross-peaks providing correlations between nuclei and thereby revealing structural information and connectivity within the molecule. There are many different 2D NMR techniques, but in this thesis the HSQC (Heteronuclear Single Quantum Coherence), HMBC (Heteronuclear Multiple Bond Correlation) and 2D INADEQUATE (Incredible Natural Abundance DoublE QUAntum Transfer Experiment) NMR sequences were used. HSQC can be used to identify one-bond correlations between ^{1}H and ^{13}C, HMBC to identify multiple-bond correlations between ^{1}H and ^{13}C, and INADEQUATE is used to identify one-bond ^{13}C-^{13}C correlations. INADEQUATE has a very low sensitivity due to the low natural abundance of ^{13}C, which makes the probability of two adjacent carbon sites both being ^{13}C very low [62].

Diffusion NMR can be used to determine the self-diffusion coefficient, D, of a compound, which gives an indication of its molecular size. In the experiment, field gradient pulses are used to encode the positions of the molecules. A second gradient pulse is then used to reverse the encoding. For molecules that have changed position due to diffusion during the time between the gradient pulses, the accumulated phase is not fully refocused, resulting in a loss of phase coherence and attenuation of the detected signal. The faster a molecule diffuses, the further it will move from its original position, and consequently, the greater the signal attenuation. The experiment is repeated for different gradient strengths, and then the diffusion coefficient can be determined by a regression fit of the signal intensity as a function of gradient strength [63].

To decipher the complexity of PG materials, both ^{1}H and ^{13}C NMR spectroscopy were used to quantitatively determine the various types of substructures that can form in the material. DEPT-135, HSQC, HMBC and INADEQUATE were used to aid the assignment of the ^{1}H and ^{13}C signals. A high concentration of PG (180 mg sample in 600 µl solvent) was used in the HSQC, HMBC and INADEQUATE experiments in **Paper I** to increase sensitivity, while 60 mg sample in 600 µl solvent was used for ^{1}H, ^{13}C and DEPT-135 experiments. Dimethyl sulfoxide-d6 (DMSO-d6) was used as solvent to enable observation of the hydroxyl protons, and experiments were conducted on an 800 MHz Bruker Avance III HD spectrometer (Oxford magnet).

NMR was also used to confirm the overall structures of the PG ester surfactants. This was done with ^{1}H NMR on Bruker Avance NEO spectrometers operating at ^{1}H frequencies of 400 MHz and 600 MHz in **Paper II** and **Paper III**, respectively, in methanol-d4 solvent to prevent the overlap of the water peak with PG peaks that can occur in DMSO-d6.

For the diffusion NMR experiments, the integral intensities of certain known signals were plotted on a logarithmic scale against $k = (\gamma g \delta)^2(\Delta - \delta/3)$, where γ is the gyromagnetic ratio of the nucleus studied, g is the gradient strength (ramped in steps from 1.01-49.32 G/cm), δ is the gradient pulse length (4 ms) and Δ is the diffusion time (100 ms). For a species characterized by a single self-diffusion coefficient, the signal attenuation is monoexponential according to

$I(k) = I_0 e^{-k}D$, where $I(k)$ is the NMR signal intensity at a certain k, I_0 is the intensity at zero applied gradient strength, and D is the self-diffusion coefficient. Heterogeneous mixtures, however, can show non-monoexponential behavior, in which case a distribution of the self-diffusion coefficients can be estimated according to the gamma distribution model [64], which can provide an indication about the average size and polydispersity of the material. Experiments were done in DMSO-d6 solvent to minimize effects of self-assembly, on an 800 MHz Bruker Avance III HD NMR spectrometer, with a concentration of 7-8 mg sample in 600 µl solvent for the esters, and similar molar concentrations for the PG3 and decanoic acid raw materials.

3.3 Surface tension measurements with the pendant drop technique

In this work, the pendant drop technique was used to measure the surface tension of the surfactant solutions. The adsorption isotherms were then used to determine the CMC of the surfactants.

In the pendant drop technique, a drop is suspended from a needle and the shape of the droplet is analyzed to determine the surface tension. The shape of the droplet is governed by a balance of the forces resulting from the gravity and the surface tension, where the gravity pulls the drop downwards and deforms it, while the surface tension counteracts this, trying to minimize the surface area and striving for a spherical drop shape. An example of this can be seen in Figure *7*, which shows two hanging droplets of equal volume (5 µl): one of pure water (Figure 7a) and one of a surfactant solution (Figure 7b). Water has a high surface tension (72 mN/m) which pulls the liquid together, making the droplet nearly spherical. For the surfactant solution however, the surfactant molecules adsorb at the interface and lowers the surface tension, allowing gravity to deform the droplet considerably more.

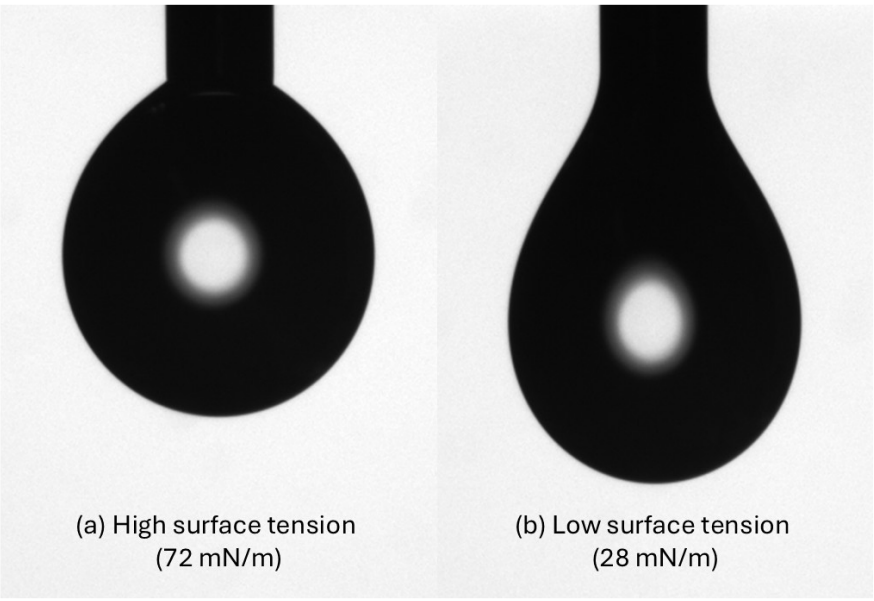

Figure 7. Two droplets of equal volume (5 µl): (a) water, with a surface tension of ca. 72 mN/m, and (b) a surfactant solution, with a surface tension of 28 mN/m. The high surface tension of water keeps the droplet close to spherical, while the lower surface tension of the surfactant solution allows the droplet to deform more.

The Young-Laplace equation can be used to explain the relationship between the surface tension and drop shape with the interfacial pressure [65]:

$$\gamma\left(\frac{1}{R_1}+\frac{1}{R_2}\right)=\Delta P \qquad (3)$$

Where γ is the surface tension, R1 and R2 are the radii of the droplet (R1=R2 for a spherical droplet), and ΔP is the Laplace pressure [65].

The surface tension was measured at several concentrations, and used to construct the surface tension isotherm, from which the CMC was estimated as the concentration of the inflection point in the curve, and the surface tension at the plateau as an average of the surface tension of the points at the plateau. In cases with a slanted plateau, the lowest surface tension was taken as the plateau value.

In **Paper III**, the surface tension was determined after 30 seconds of equilibration time, to allow the surfactants to diffuse to the interface, while keeping the measurement time not too long to prevent effect from an increase in concentration due to water evaporation. In **Paper II** however, the surface tension was also studied as a function of time. To be able to measure the surface tension over longer times (up to 360 seconds in this case) without having a considerable effect from water evaporation, the measurements were conducted inside a quartz cuvette with water, to have a humid environment and minimize evaporation. The setup can be seen in Figure 8.

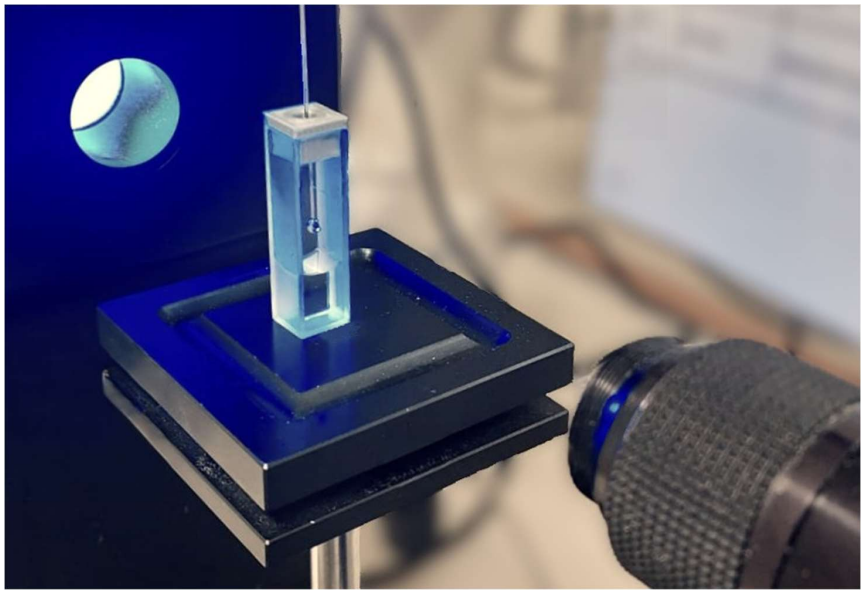

Figure 8. Set up for the surface tension measurements in a confined, high humidity environment. At the top left is a light source. In the middle is a droplet suspended from a needle inside a quartz cuvette. The cuvette has water at the bottom and is covered with a lid. At the bottom right is the camera recording images of the drop shape.

3.4 Turbidity estimation of surfactant solutions with UV-vis spectroscopy

Micellar solutions of surfactants in water are typically transparent, since spherical micelles, with radii typically in the order of a few nanometers [9], are small enough that they scatter visible light to little extent. Larger surfactant aggregates, however, can scatter light to larger extent, leading to a turbid appearance. As was mentioned in the background section, solutions of ethoxylated surfactants can turn cloudy upon heating as a conformational change in the polyoxyethylene head group promotes formation of larger self-assemblies that scatter light. Whether the same phenomenon occurs for polyglycerol surfactants is unclear. Clouding has been reported for PG surfactants in the literature [29, 66-68], but whether it arises from a conformational change similar to that of ethoxylated surfactants remains uncertain.

To study the temperature dependence of the solution behavior of the surfactants used in this thesis, UV-vis spectroscopy was employed. In UV-vis spectroscopy, a light source with well-defined wavelengths shines on a sample with intensity I_0, and the transmitted light (I) is measured by a detector. The absorbance (A) is then calculated as $A = \log_{10}(I_0/I)$ [69]. For turbid dispersions, the larger aggregates scatter light, so the intensity of the transmitted light is lower, giving a higher measured absorbance. Absorbance can therefore be used as a measure of the turbidity of the solution.

In this thesis, absorbance was measured at a wavelength of 400 nm using a Hewlett Packard 8453 UV-vis spectrophotometer. Measurements were performed on 1 wt% solutions at temperatures between 16 and 56 °C.

3.5 Double syringe foam test

To test the foamability and foam stability of surfactants, a modified version of the double syringe method [70] was used. Two plastic syringes were connected via a short plastic tube, where one syringe is empty, and one contains half the volume of surfactant solution and the rest air, see Figure 9 for the experimental setup. The content of the syringe is then pushed through the narrow connection to the next syringe 20 times, creating a foam. Then the content is quickly transferred to a measuring cylinder to observe the initial foam volume, and the development over time. In **Paper II**, 2 ml Braun Injekt® Luer Solo plastic syringes were used, and they were filled with 1.5 ml of a 0.05 wt% surfactant solution (approximately 2 × CMC) and the rest air (ca 1.5 ml air), and the foam and liquid were then transferred to a 5 ml measuring cylinder.

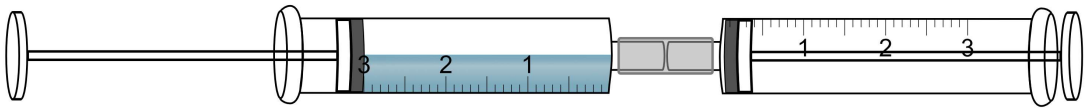

Figure 9. Setup for the double syringe method. Two plastic syringes were connected via a short plastic tube, where one contains surfactant solution and air. The content is then pushed back and forth between the two syringes to create a foam.

3.6 Dynamic light scattering

Dynamic light scattering (DLS) is a technique used to measure the size distribution of particles in a liquid, utilizing the fact that smaller particles diffuse faster than larger ones. The sample is irradiated by light, which is scattered by the particles, and as the particles diffuse due to Brownian motion, the scattering fluctuates over time. The light scattering intensity can therefore

be correlated to the diffusion of the particles. The particle size can then be calculated, as the size and diffusion coefficient are inversely proportional according to the Stokes-Einstein equation:

$$d = \frac{k_B T}{3\pi\eta D} \quad (4)$$

where D is the diffusion coefficient, k_B is the Boltzmann constant, T is the temperature, η is the viscosity and d is the hydrodynamic diameter of the particle [71].

A limitation with DLS is that it measures the intensity-weighted size distribution, and will underestimate the amount of small particles in a wide distribution of particle sizes. By making certain assumptions, for example the refractive index of the particles, the volume- and number-weighted distributions can however be estimated [71]. Another limitation is that the size determined by DLS is the hydrodynamic diameter, defined as the diameter of a sphere that diffuses at the same rate as the particle, and is therefore a rough approximation for particles of other shapes [72].

In this work, DLS was used to study the aggregate size of surfactant self-assemblies. For spherical micelles, the radius of the hydrocarbon core is expected to be close to the extended length of the surfactant's hydrophobic tail, ca. 1.5-3.0 nm [9], with the full micelle radius somewhat larger including the hydrophilic head. Other self-assembled structures can be considerably larger, and they can be far from spherical, meaning the hydrodynamic diameter measured for these is an approximation. The DLS measurements were done on a Litesizer 500 instrument (Anton Paar), with surfactant concentrations of approximately 1.5 × the surfactant CMC. Volume-weighted size distributions were estimated by assuming a refractive index of 1.45.

3.7 Quartz crystal microbalance with dissipation monitoring

The quartz crystal microbalance with dissipation monitoring (QCM-D) technique is, as the name suggests, a gravimetric way to measure very small masses adsorbed on a quartz crystal surface. It uses the piezoelectric properties of the quartz crystal, which relates its resonance frequency to the mass of the crystal. A small change of mass directly affects the resonance frequency [73]. Therefore, as molecules are adsorbing to the surface, the mass of the crystal changes, and the resonance frequency changes. The Sauerbrey equation [74] can be used to determine the adsorbed mass from the change in frequency:

$$\Delta m = -\frac{\Delta f \cdot C}{n} \quad (5)$$

where Δm is the adsorbed mass per unit area, Δf is the frequency shift at overtone n and C is an equipment characteristic constant (17.7 ng·Hz^{-1}·cm^{-2} for a 5 MHz crystal) [75, 76].

To study the adsorption on various surfaces, the surface of the sensors can be modified to provide various surface chemistries. Gold coated QCM-D sensors can for example be modified with various thiols with different terminal groups, forming self-assembled monolayers (SAMs) on the surface [77]. Thiols with hydroxyls at the end group give a hydrophilic functionalization, while a methyl group at the end gives a hydrophobic functionalization.

The adsorbed mass determined with QCM-D includes not only the adsorbed molecule, but also trapped solvent and the hydration layer, and therefore represents the hydrated mass for surfactant in water solution. For surfactant adsorption, typical adsorbed masses measured with QCM-D are 150-250 ng/cm^2 at hydrophobic surfaces, representing a monolayer or hemimicelles where the surfactant tails are turned towards the hydrophobic surface [78-80]. At hydrophilic

surfaces, the surfactants tend to form bilayers on the surface instead, as the polar heads are turned both towards the surface and the solution. This typically yields an adsorbed mass up to double that of the adsorbed monolayer for a full and dense bilayer, or lower in case of interdigitated tails [78].

In addition to measuring the frequency change, the change in dissipation (ΔD) is also measured. A high ΔD indicates a soft and swollen adsorbed layer [81]. When the ΔD is high ($>0.5 \cdot 10^{-6}$) the Sauerbrey equation is not valid as it assumes a thin, rigid adsorbed layer [76]. Although the adsorbed mass cannot be accurately calculated with the Sauerbrey equation in these cases, the high shift in dissipation itself gives valuable information about the nature of the adsorbed layer.

In this work, gold coated QCM-D sensors (5 MHz, 14 mm Cr/Au QCM-D sensors from QuartzPro) were modified with 1-hexadecanethiol or 16-mercaptohexadecan-1-ol to form hydrophobic and hydrophilic SAMs on the surface, respectively. Before modification, the sensors were cleaned by 10 min UV-ozone treatment, 5 minutes immersion in a 5:1:1 (v:v) mixture of Milli-Q water, ammonia (25 %) and hydrogen peroxide (30 % w/v) at 75 °C, rinsing with Milli-Q water and drying with nitrogen gas, followed by another 10 min of UV-ozone treatment. The gold surfaces were immersed in a thiol solution of 2 mM thiol in ethanol overnight to form the SAM functionalization. After surface modification, the surfaces were rinsed with first ethanol and then water to remove any remaining thiols that had not attached, and was then dried with nitrogen gas.

For the QCM-D measurements, the modified sensors were placed in the instrument and first Milli-Q water was pumped through the flow cell to record a baseline. Then, surfactant solutions of increasing concentrations were pumped through the flow cell one at a time at a rate of 0.100 ml/min. For each concentration, Δf was monitored, and as it stabilized at a plateau, the next concentration would be added. Δf of the third overtone was used for determination of the adsorbed mass with the Sauerbrey equation.

3.8 Preparation of AuNP spectra

To evaluate the ability of CNN models to predict the size distributions of AuNPs, training and predictions were made on simulated spectral data. ML, and in particular deep learning algorithms such as CNNs, require large training datasets, which is challenging to obtain with experimental data only. Therefore, the work here was conducted only on simulated data, as a proof of concept.

The strategy taken was through supervised learning with CNN models, that would take a simulated extinction spectrum of a colloidal gold particle mixture as input, and its respective parameters for the size distribution of the particles as output. The model would then train on these input-output pairs to learn the correlation between the appearance of the spectrum with the size distribution. To create these input-output pairs, simulated extinction spectra of single gold nanoparticles of certain sizes were used. These single-particle spectra (simulated using Finite-Difference Time-Domain simulations, which is out of scope for this thesis) were obtained for spherical AuNPs with diameters in the range of 10-100 nm, with a step size of 2 nm, and rod-shaped AuNPs with these same diameters, and aspect ratio (AR) in the range 1-4 with a step size of 0.25. All spectra were simulated for AuNPs in water environment. Visual inspection of the spectra was made to check for errors in the simulation, and deviations in the data were replaced by interpolations. The wavelengths of the spectra were limited to 400-800 nm for

spherical particles, and 400-1000 nm for rods, with inclusions of the longer wavelengths for rods as they contain more information at high wavelengths compared to the spectra for spheres.

To create spectra for polydisperse mixtures of gold nanoparticles of a certain size distribution, single particle spectra were summed together according to Gaussian size distributions. For each spectrum, the mean and standard deviation of the size distribution was randomized, and 100k particle sizes were drawn from this distribution and added together to form a spectrum representing the mixture of particles. For the rod-shaped particles, two distributions were randomized: one for the diameter and one for the aspect ratio. The resulting spectra were then min-max scaled to have extinctions between 0-1. Datasets of 1500 and 3000 spectra were generated for spherical and rod-shaped particles respectively.

To test the robustness of the model, some datasets were also prepared with different levels of noise in the spectra. A value drawn from a normal distribution with mean 0 and standard deviation of 0.0025, 0.0050, 0.0100, 0.0200, or 0.0300 was added to each datapoint in the spectrum, where a larger standard deviation gives a noisier spectrum. The spectra were then min-max scaled again.

3.9 Construction of convolutional neural networks for size distribution predictions

For the task of determining the heterogeneity of gold nanoparticle mixtures from their extinction spectra, the convolutional neural network (CNN) architecture was chosen, as CNNs are well established for recognizing patterns in image-like data. The reasoning was that, in the same way a CNN model can learn to recognize the pattern of a dog or a house in a two-dimensional image, it should be able to learn the characteristic features of an AuNP extinction spectrum, although in one dimension instead of two.

The CNN models were constructed in Python (version 3.8.3), using the Keras [82] deep learning framework (version 2.5.0). The model architecture and hyperparameters were tuned to keep the training time low while maintaining good performance in terms of the mean squared error (MSE) and the coefficient of determination (R^2). The resulting architectures of the two models (for spherical and rod-shaped particles, respectively) are shown in Table 1. Although the two-dimensional (2D) versions of the convolutional and max-pooling layers are used to make them compatible with other functionalities (out of scope for this thesis), the size of the first dimension is 1, effectively making the layers one-dimensional and suitable for the 1D spectral input. The number of neurons in the final layer (2 and 4, respectively) corresponds to the parameters of the predicted size distributions, i.e. the mean and standard deviation of the diameter distribution for the spheres, and diameter and AR distribution for the rods.

ReLU was used as the activation function for the convolutional layers and the first fully connected (dense) layer. The output layer (the second dense layer) had no activation function, as is standard for regression tasks. The dataset was split into 70% training and 30% test data, and 30% of the training data was used for validation during training. The Adam optimization algorithm was used with a learning rate of 0.001. Training was performed with a batch size of 100 for 600 epochs, using MSE as the loss function, with the ModelCheckpoint callback used to retain the weights from the epoch with the lowest validation loss. Three replicates were made for each level of noise on the data, by using different TensorFlow random seeds.

Table 1. Architectures of the two CNN models used for prediction of AuNP size distribution. Recreated with permission from [83].

Layer	**Parameters for the sphere model**	**Parameters for the rod model**
Conv2D	Filters: 8 Kernel size: (1,150)	Filters: 16 Kernel size: (1,50)
Conv2D	Filters: 16 Kernel size: (1,120)	Filters: 32 Kernel size: (1,40)
MaxPooling2D	Pool size: (1,2) Strides: (1,2)	Pool size: (1,2) Strides: (1,2)
Conv2D	Filters: 32 Kernel size: (1,80)	Filters: 64 Kernel size: (1,20)
Conv2D	Filters: 64 Kernel size: (1,40)	Filters: 128 Kernel size: (1,10)
MaxPooling2D	Pool size: (1,2) Strides: (1,2)	Pool size: (1,2) Strides: (1,2)
Dense	Neurons: 30	Neurons: 60
Dense	Neurons: 2	Neurons: 4

4 Results and discussion

4.1 Heterogeneity of the polyglycerol building blocks (Paper I)

As was discussed in the introduction and background sections, PG constitutes the hydrophilic building block in PG ester surfactants. The properties of PG surfactants depend to a large extent of the size distribution of the PG oligomers, as it will affect the HLB. One might also expect the molecular architecture of the PG, in terms of linear or branched structures, to have an effect on the properties of the surfactants. The PG structure has been shown to affect its properties in other fields [84-88], so it can be assumed that it would also have an impact when used for preparing surfactants. Whether the PG is branched or linear could, for example, potentially affect the shape of the surfactant molecule, and in turn the CPP and its self-assembly structures. Characterization of the heterogeneity present in the PG raw material is therefore an important first step toward understanding structure-property relationships in the final surfactant systems. In **Paper I**, the aim was therefore to develop a method for quantification of this structural heterogeneity in PG.

The heterogeneity of PG arises from the multiple possible ether bonds that can form during glycerol condensation, resulting in mixtures of molecules with different degree of polymerization and constitutional arrangements. The structure of the glycerol units in PG has been described previously in literature using the broad categories glycerol (G), terminal (T), linear (L) and dendritic (D) units, according to the number of connections the glycerol unit has to other glycerol units (zero, one, two and three, respectively) [89-91]. Some groups also differentiate between the two types of T and L structures that arise from whether the connections are via the primary or secondary carbons. Here we have chosen to denote them Tp, Ts, Lpp and Lsp, where the p and s denote connections via the primary or secondary carbons, respectively. In the present work, this description was extended further by considering whether the ether linkages involved primary or secondary carbon atoms on both of the two connecting glycerol units for each ether bond. This extended categorization provides a more detailed picture of the constitutional heterogeneity present in the samples, allowing discrimination between substructures that would otherwise be grouped together. In total, 18 distinct substructures can be identified in this way, illustrated in Table 2.

Table 2. The 18 possible substructures of polyglycerol, classified according to the number of connections a glycerol unit has to other glycerol units (G, T, L or D), and further divided according to whether the ether linkages involve primary (p) or secondary (s) carbon atoms. In the substructure labels, letters before the hyphen denote the carbon(s) of the glycerol unit in question, and letters after the hyphen denote the corresponding carbon(s) on the neighboring glycerol unit(s), with letters paired in order across the hyphen.

Structure	G	Tp	Ts	Lpp	Lsp	D
Substructure	G	Tp-p Tp-s	Ts-p Ts-s	Lpp-pp Lpp-ps Lpp-ss	Lsp-pp Lsp-ps Lsp-sp Lsp-ss	Dpsp-ppp Dpsp-pps Dpsp-psp Dpsp-sps Dpsp-pss Dpsp-sss

To investigate whether the heterogeneity of PG could be quantified in terms of these 18 substructures, four commercial PG samples with increasing average degrees of oligomerization (PG2, PG3, PG4 and PG6) were analyzed using a combination of one- and two-dimensional NMR experiments. Although quantitative compositional analysis is commonly performed using ^{1}H NMR spectroscopy, quantitative ^{13}C NMR was chosen as the main analysis method in this work due to the high structural complexity of polyglycerol which results in heavily overlapping signals in ^{1}H NMR. The ^{13}C NMR spectra of the four samples can be seen in Figure 10.

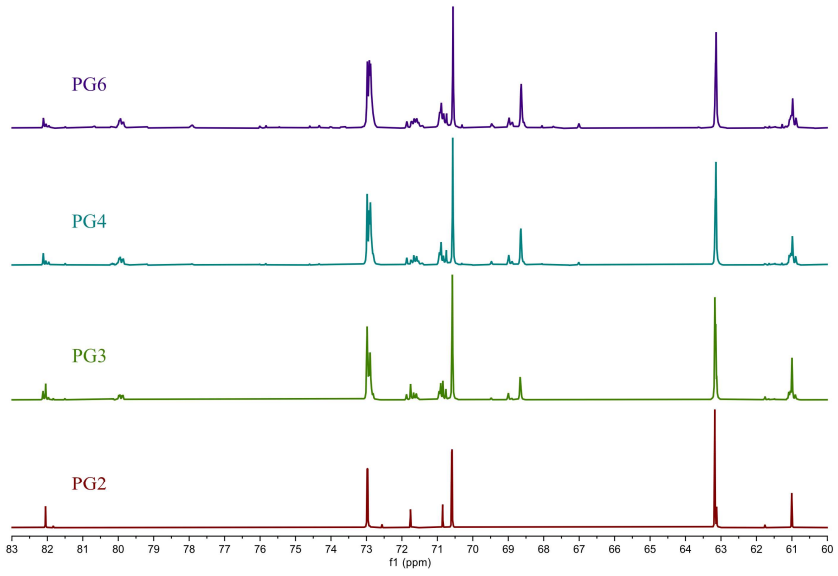

Figure 10. ^{13}C NMR spectra of the four PG samples recorded in DMSO-d6 at 201 MHz.

4.1.1 ^{13}C NMR peak assignment

The first step in the attempt to quantify the heterogeneity of PG from its ^{13}C NMR spectrum, was to assign the carbon signals in the spectra to the substructures in Table 2. The assignments were based on the samples PG2 and PG6, representing the smallest and largest average degree of polymerization, respectively, in combination with pure glycerol, as these three samples together covered all observed PG carbon signals. By combining quantitative ^{13}C NMR with DEPT-135, INADEQUATE, HSQC and HMBC experiments, the majority of the carbon signals could be assigned to one or more of the 18 glycerol substructures, see Figure 11. Not all signals could be assigned however, and some peaks with overlapping signals were assigned to several glycerol substructures. It is also worth noting that only seven of the 18 substructures were identified within the four samples, as will be discussed further in the next subsection.

The assignments were compared to assignments from previously published studies [89, 90, 92], see Figure 12. The comparison with the work by Cassel et al. [92] is especially relevant, as it also contains information on a substructure level. They report the chemical shifts for well-defined oligomer standards of various PG molecules, which is methodologically very different from our work in **Paper I**, but the assignments show overall good agreement, and can be regarded as validation of our assignments. The chemical shifts of some of their oligomer standards give valuable information about glycerol substructures that were not identified within our samples, and can suggest the identity of some of our unassigned peaks. In contrast, our methodology presents a way to determine substructure information from complex mixtures, without the need for comparison with standards. While not on a substructure-level, the assignments done by Sunder et al. [89] have been widely adopted for characterization of PG structure, and have since been further expanded, as is shown in [90]. The material in Sunder's work is hyperbranched PG with much higher degree of polymerization than in **Paper I**, so the spectral response is quite different with a high degree of Lsp and D structures. Regardless, the assignments in our work and those by Cassel et al. suggest that the underlying complexity of the material could be described in more detail by assignments on a substructure level.

The assignment methodology in **Paper I** enabled identification not only of glycerol units according to their connectivity (glycerol, terminal, linear or dendritic), but also according to the type of neighboring unit and the nature of the linkage connecting them. As a result, the constitutional heterogeneity of the PG samples could be characterized at a significantly higher level of detail than is possible using conventional assignment approaches.

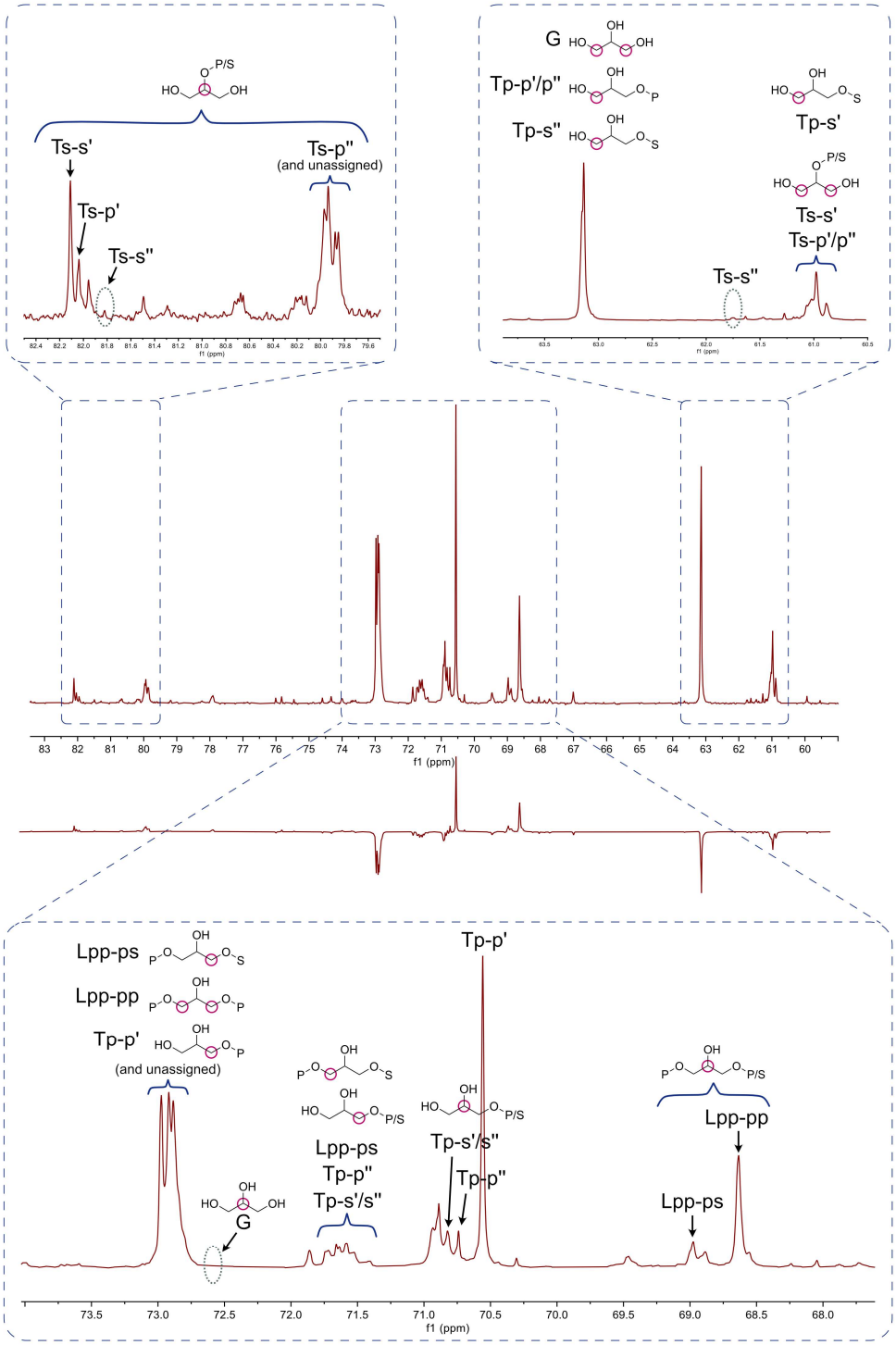

Figure 11. ^{13}C NMR spectrum of PG6, with the signal assignments shown. In cases where an assigned signal is missing or below the signal-to-noise threshold, the position is shown with a dashed circle. In cases where a specific substructure was found for two sets of signals, these are marked with prime (′) and double prime (″) in the figure. The corresponding DEPT-135 spectrum is shown right below the ^{13}C NMR spectrum, showing CH (secondary carbons) as positive peaks and CH_2 (primary carbons) as negative peaks. Figure from Paper I.

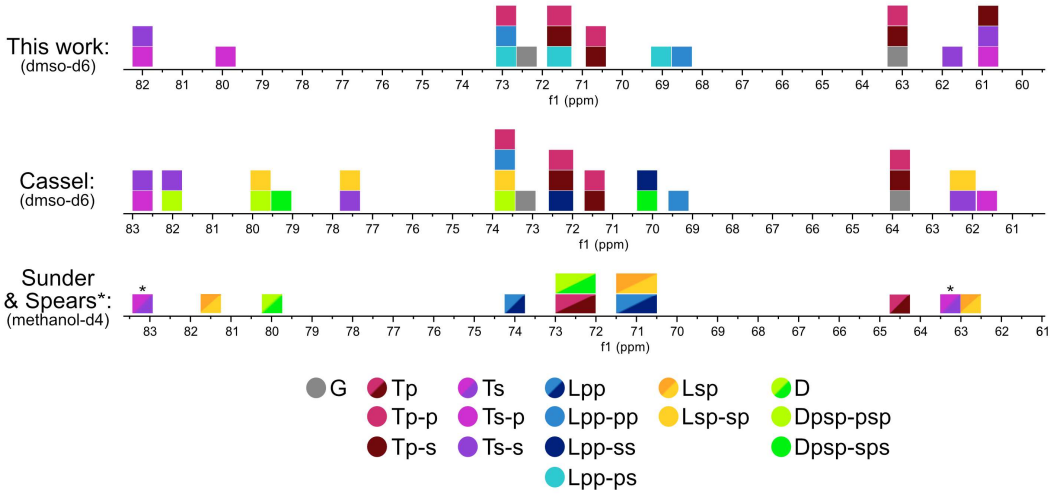

Figure 12. Comparison of the peak assignments from **Paper I** to assignments from Cassel et al. [92] and Sunder et al. [89] complemented with assignments from Spears et al. [90]. The spectra were recorded in different solvents and at different concentrations, which affects the chemical shifts, and the axes have therefore been adjusted to facilitate visual comparison of the assignments. Figure from Paper I.

4.1.2 Structural diversity revealed by NMR spectroscopy

The ^{13}C NMR spectra of the PG samples, see Figure 10, showed an increase in spectral complexity with increasing oligomer size. While PG2 exhibited a relatively small number of well-defined peaks, PG6 displayed a substantially larger number of signals, reflecting the growing number of possible constitutional arrangements as the degree of polymerization increases. Most of the observable carbon signals could be assigned to specific glycerol substructures. As is described in detail in **Paper I**, deconvolution was then used to get the integral of each signal, to quantify the relative amounts of each substructure, see Table 3 for the results. Seven of the 18 possible substructures were identified in the samples. An important observation was the absence of signals originating from dendritic structures in the four samples. This is consistent with supplier information and indicates that the investigated materials are predominantly linear oligomers. Lsp structures were not assigned either. Nevertheless, the increasing number of unassigned peaks with increasing oligomer size suggests that these or other of the possible substructures are present, although in lower relative amounts. The unassigned contribution increased from approximately 1% in PG2 to 22% in PG6, illustrating how the number of possible connections and the probability of finding them expands as the molecules become larger.

Table 3. Relative amounts of each identified substructure in the four PG samples, based on integrals in the quantitative ^{13}C NMR spectra.

Sample	G	Tp-p	Tp-s	Ts-p	Ts-s	Lpp-pp	Lpp-ps	Unassigned
PG2	2%	72%	12%	11%	1%	1%	-	1%
PG3	-	55%	6%	10%	3%	14%	4%	8%
PG4	-	43%	4%	8%	3%	21%	5%	15%
PG6	-	35%	8%	5%	3%	22%	5%	22%

As expected, the fraction of terminal units decreased with increasing oligomer size, while the fraction of linear units increased. A key finding was the predominance of primary ether linkages for all samples. Among the identified substructures, Tp-p and Lpp-pp were consistently the most abundant of the T and L structures, respectively. This demonstrates that glycerol condensation preferentially occurs through primary hydroxyl groups, resulting in mainly linear growth of the oligomer chains, which is supported by supplier claims [93] and literature [24, 25].

For the PG2 sample, the degree of polymerization provided by the supplier indicates that it is constituted mainly by diglycerol, and this can help in establishing the molecular structures based on the relative abundance of the substructures. 1% is linear Lpp-pp, and assuming that there are no PG molecules with more than three glycerol units, these Lpp-pp structures are the center of a triglycerol molecule, connected to Tp-p substructures on each side. 2% of the glycerol units belong to unreacted glycerol, and as no other linear or dendritic substructures were found, the remaining terminal structures can be assumed to form diglycerol molecules in pairs. The remaining Tp-p structures can only form diglycerol with other Tp-p structures, as the neighbor also has to be connected via its primary carbon. Similarly, the Ts-s structures must form dimers with themselves. Finally, the Tp-s and Ts-p structures must form diglycerol together, and as can be seen from Table 3, their relative amounts of 12 and 11% respectively match this very well. This results in a sample composition that is presented in Figure 13. For the other samples, the distribution of degree of polymerization prevents the same analysis to be conducted, which further demonstrates the challenge of characterizing heterogeneous materials.

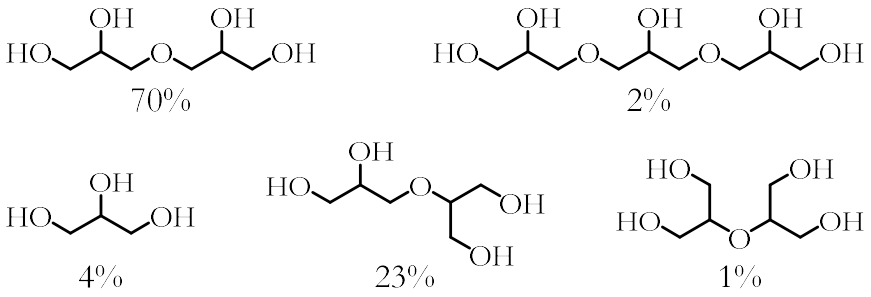

Figure 13. Molecular composition of the PG2 sample, based on the ^{13}C NMR analysis, reported as the mol% of each molecular structure in the sample mixture.

Although the exact molecular composition cannot be established for the larger PG samples in the same way as for PG2, the quantification of the various substructures still gives valuable information. It gives more detailed information about the structure than what was previously attainable with the more general categorization found in literature. This can be useful as a benchmark when changing supplier or synthesis method for PG materials, or as a tool for establishing structure-property relationships.

For a more comprehensive quantification, future work could involve using this methodology on more diverse samples, containing a considerable amount of D and Lsp structures, to identify their chemical shift.

4.1.3 Simplified assessment of heterogeneity using ^{1}H NMR

While quantitative ^{13}C NMR provides extensive constitutional information, it is experimentally demanding. Therefore, a complementary approach based on quantitative ^{1}H NMR was also developed.

Although the ^{1}H NMR spectra were difficult to interpret due to the many overlapping signals, the hydroxyl region of the spectrum could be assigned to G, T or L structures by using the assignments of the ^{13}C spectra and HSQC 2D spectra, see **Paper I** for details about the assignment. Analysis of the hydroxyl proton region enabled an estimation of the relative fractions of glycerol, terminal and linear units, although information on the detailed substructure-level was not possible. The results obtained from ^{1}H NMR, see Table 4, were in good agreement with those derived from quantitative ^{13}C NMR, although not giving as detailed information. It should be noted that the L structure mentioned here is not necessarily covering all possible L structures, as only Lpp-pp and Lpp-ps were identified in the four samples.

Table 4. Relative amounts of G, T, L and unassigned structures in the four PG samples determined from ^{1}H (left), with a comparison with the results obtained from ^{13}C NMR (right).

	Relative amounts from ^{1}H NMR				Relative amounts from ^{13}C NMR			
Sample	G	T	L	Unassigned	G	T	L	Unassigned
PG2	2%	93%	3%	2%	2%	96%	1%	1%
PG3	-	73%	20%	7%	-	74%	18%	8%
PG4	-	58%	26%	16%	-	58%	26%	15%
PG6	-	48%	27%	25%	-	51%	27%	22%

The ^{1}H NMR method provides a practical tool for rapid characterization of PG heterogeneity when detailed substructural information is not required. This is particularly useful for routine analysis of PG raw materials intended for surfactant synthesis, where the balance between glycerol, terminal, linear and unassigned units may already provide valuable information regarding expected surfactant headgroup architecture.

4.1.4 Implications for surfactant heterogeneity

The results confirm that commercial PG cannot be regarded as single molecular species, but instead consist of distributions of constitutionally distinct oligomers. Furthermore, the degree of constitutional heterogeneity increases with increasing oligomer size. This type of characterization can be useful when considering switching between PG products or suppliers, or for routine control of PG batches after synthesis, to ensure consistent material characteristics.

For surfactant systems, which will be discussed in the next section of this thesis, the characterization of the PG raw material is highly relevant. To establish structure-property relationships, one must first know the structure of the building blocks. The constitution of the polar head group could, for example, affect the CPP of the surfactant, which in turn is expected to affect the packing at interfaces and the self-assembly, and the physicochemical properties of the product. It is therefore of interest to have a straightforward NMR methodology to quantify the structure. In addition, the position of the hydrophobic tail on the PG head group could also affect the structure, and in turn the properties, of the surfactants. Although not studied here, the methodology developed in this work could potentially be extended to also identify the tail positions, and this is a potential next step in determining the quantitative structure-property relations of PG ester surfactants.

4.2 Heterogeneity in polyglycerol ester surfactants (Paper II and III)

The previous section demonstrated that commercial PG materials are inherently heterogeneous, consisting of distributions of oligomers with different chain lengths and constitutional structures. During surfactant synthesis, this complexity is further amplified, as was discussed in the introduction section. Esterification can occur at different hydroxyl groups within the PG molecule and with varying degrees of substitution, resulting in distributions not only in headgroup structure but also in the number and position of attached hydrophobic chains.

Consequently, polyglycerol ester surfactants are highly heterogeneous systems. This raises two important questions. First, how does this heterogeneity influence the surfactant behavior? Second, can the underlying structure-property relationships be revealed by studying well-defined model systems in which the effects of molecular heterogeneity are minimized? These questions were addressed through two complementary studies. In the first study (**Paper II**), the effect of heterogeneity in degree of esterification was investigated using PG3C10 surfactants with different monoester contents. In the second study (**Paper III**), structurally well-defined model surfactants were synthesized to study the influence of molecular architecture without the complication of compositional heterogeneity.

4.2.1 Influence of esterification heterogeneity on surfactant properties (Paper II)

To investigate the effect of heterogeneity in terms of degree of esterification, three PG3C10 samples with different compositions were prepared. One sample was synthesized using a PG-to-fatty acid molar ratio of 2:1 (sample 2:1), one using a ratio of 5:1 (sample 5:1), and a third sample was purified to yield almost exclusively monoesters (sample ME). The monoester content was therefore expected to increase in the order 2:1 < 5:1 < ME.

The compositional differences between the samples were confirmed using diffusion NMR experiments, in DMSO-d6 solvent to prevent self-assembly. As the PG3 raw material and the three surfactant samples are heterogeneous mixture, the attenuation data could not be used to find one single diffusion coefficient, but instead the data was fitted according to a gamma distribution model, for which the mean diffusion coefficient and width of the distribution are reported in Table 5. Decanoic acid could however be assigned a single diffusion coefficient. The average diffusion coefficient increased from the 2:1 sample to the ME sample, which is consistent with a decreasing proportion of larger and more hydrophobic multiester species, as a smaller molecule diffuses faster. The results demonstrate that the three materials represent surfactant systems with progressively decreasing average degree of esterification while retaining the same PG headgroup size and fatty acid tail length.

Table 5. Mean diffusion coefficients (Dmean) and distribution widths (σD) for the three samples and the PG3 and decanoic acid raw materials, obtained from fits of the diffusion NMR attenuation data. Uncertainties represent fitting errors.

Sample	Dmean ($\times 10^{-11}$ m^2 s^{-1})	σD ($\times 10^{-11}$ m^2 s^{-1})
2:1	17.78 ± 0.02	3.98 ± 0.06
5:1	18.67 ± 0.02	3.55 ± 0.08
ME	19.61 ± 0.13	3.68 ± 0.43
PG3	20.33 ± 0.10	2.91 ± 0.42
Decanoic acid	31.93 ± 0.12	–

4.2.1.1 Effect on solution behavior

The solution behavior was strongly influenced by composition, as can be seen in Figure 14. While the ME sample produced transparent aqueous solutions throughout the investigated concentration range, the 5:1 and 2:1 samples became increasingly turbid as concentration increased. Turbidity appeared at lower concentrations for the 2:1 sample. The observed trend is consistent with the expected behavior for an increased number of di- and higher esters, as the higher esters potentially have a low water solubility, and a higher CPP which would promote formation of larger self-assembly structures that could make the solution appear turbid. Regardless of the exact mechanism, the results show that relatively small changes in composition can lead to substantial differences in solution behavior.

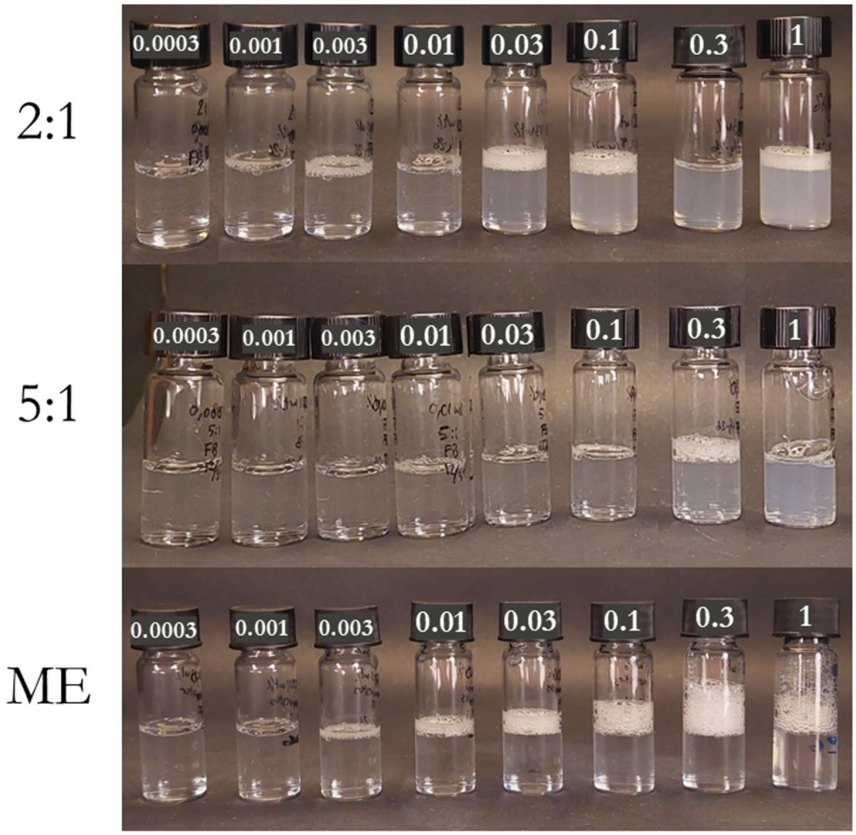

Figure 14. Aqueous surfactant solutions at room temperature (22°C) of increasing concentration for the three PG3C10 samples. The 2:1 and 5:1 samples are slightly turbid from 0.03 and 0.3 wt% and up, respectively, while the ME sample was transparent in all concentrations in the evaluated range (up to 1 wt%). Figure from Paper II.

Differences in solution behavior were also seen at elevated temperatures, see Figure 15. The 2:1 and 5:1 samples exhibited a clouding behavior, where the 1 wt% solutions became more turbid upon heating, at 55 and 50 °C, respectively, whereas the monoester-rich sample did not show such behavior in the investigated temperature range (16-56 °C). The results for ME should, however, be interpreted with caution. The 1 wt% ME solution used in this experiment was already slightly turbid at room temperature, due to partial hydrolysis of the sample at the time of this measurement. The ester bond is sensitive to hydrolysis in presence of acid or base [94], and the ME sample had some formic acid left from the flash column separation, which is suspected to have caused degradation of the sample over time.

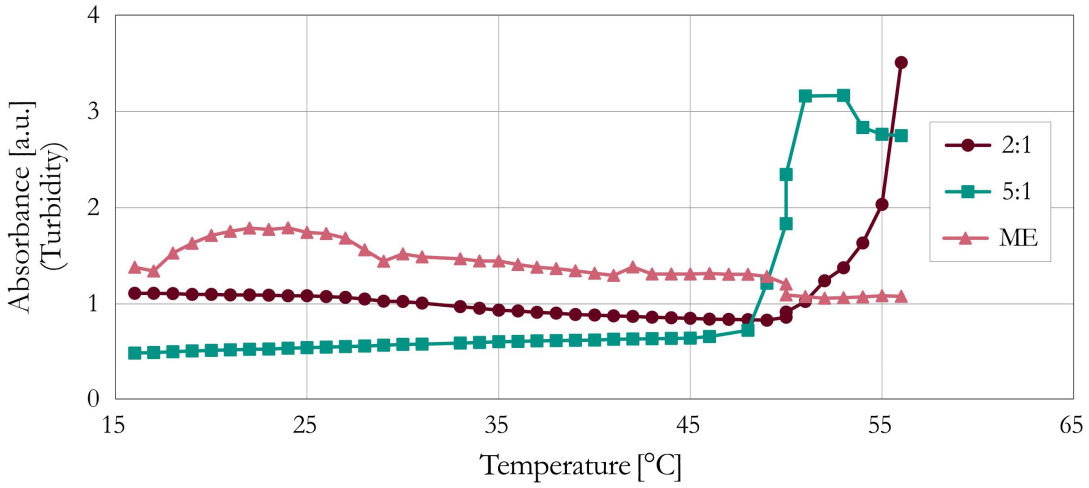

Figure 15. Turbidity at different temperatures of 1 wt% solutions of the three samples, measured as the absorbance of light at 400 nm. Figure from Paper II.

As was discussed in section 2.1.1, ethoxylated surfactants typically undergo a conformational change upon heating, yielding a less soluble headgroup, that makes the solutions turn cloudy. While clouding behavior has been reported for PG surfactants previously [29, 66-68], the mechanism behind this behavior has not been established. One hypothesis is that the PG could undergo a similar conformational change as that of polyoxyethylene, see Figure 16, as it can have a similar backbone structure when the constitution of the oligomer contains Lsp glycerol substructures, as was discussed in section 4.1. However, as was concluded based on the NMR characterization of the PG3 material, no Lsp substructures were identified, and only 8 % of the substructures are unassigned, which makes this hypothesis unlikely to be the full explanation behind the clouding behavior in PG surfactants. More work is needed to understand this phenomenon better, and to investigate what is the reason for the change in turbidity in PG surfactants.

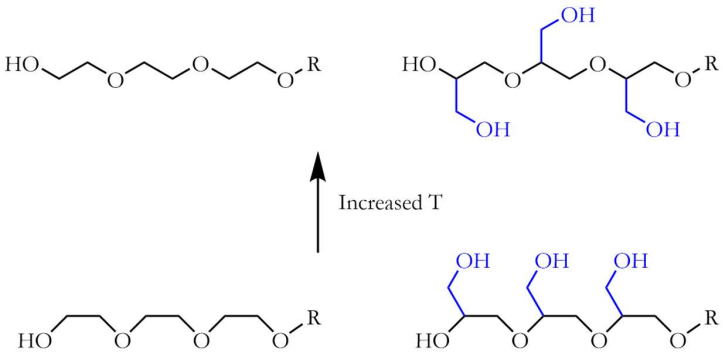

Figure 16. Illustration of the conformational change that happens in ethoxylated surfactant headgroups upon heating, to the left, and to the right, a hypothesized similar conformational change in a PG head group, if the constitution of the PG backbone is similar to that of polyoxyethylene (marked in black in the PG structure). The hydrophobic tail is denoted R.

4.2.1.2 Effect on adsorption at the air-water interface

The influence of heterogeneity was also clear when evaluating the adsorption kinetics at the air-water interface. All samples displayed relatively slow adsorption, which has been observed previously for polyglycerol ester surfactants [95]. However, the monoester-rich sample reached equilibrium significantly faster than the more heterogeneous samples, see Figure 17.

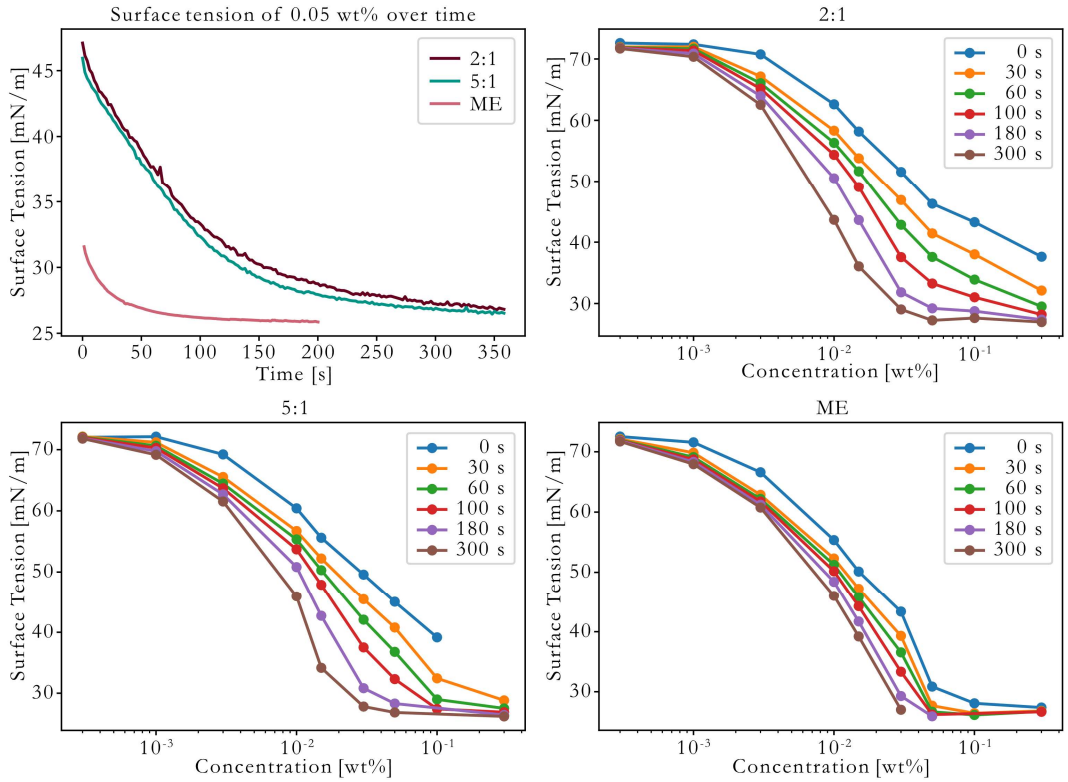

Figure 17. Surface tension at the air-water interface over time, for the three PG3C10 surfactants. The top left figure shows the surface tension as a function of time for 0.05 wt% solutions of the three surfactants, and it can be seen that the ME sample reached equilibrium faster than the other two, which are similar in adsorption kinetics. The other three figures show surface tension versus concentration, one subplot for each surfactant sample, at various equilibration times. Figure from Paper II.

Interestingly, the equilibrium properties were much less sensitive to composition. After an equilibration time of five minutes, all three samples reached similar low plateau surface tension values of approximately 26–27 mN/m and exhibited similar CMCs of approximately 0.02-0.03 wt%, see Figure 18. It should however be noted that full equilibrium was not reached after five minutes for all concentrations, and the CMC values reported here are thus not at true equilibrium. Despite this, it can be concluded that heterogeneity in terms of degree of esterification had a pronounced effect on adsorption kinetics while exerting only a minor effect on the equilibrium surface activity. This distinction is important because many practical applications involve the continuous formation of interfaces. Foaming, wetting and emulsification are examples of dynamic processes where adsorption kinetics are important for the performance.

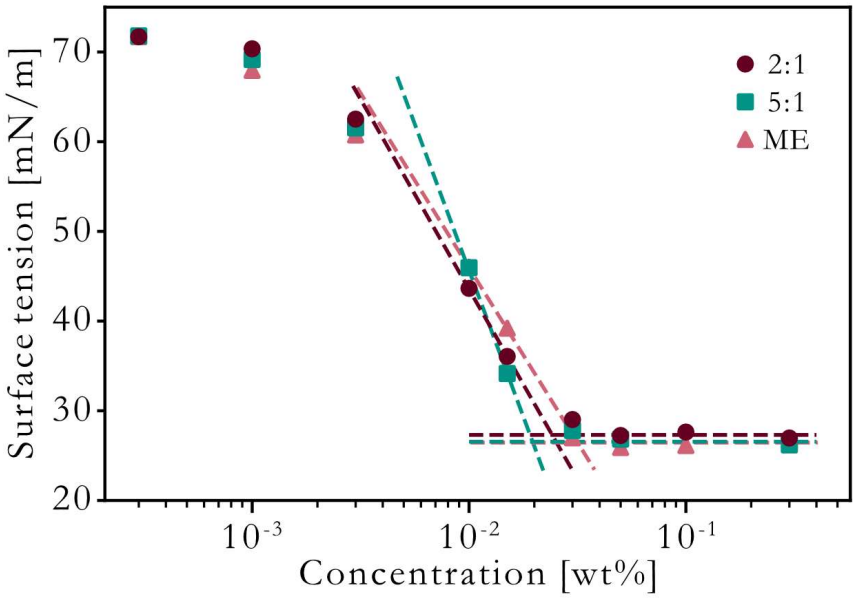

Figure 18. Surface tension isotherms for the three samples after 5 minutes of equilibration time. Figure from Paper II.

4.2.1.3 Effect on foaming performance

Foam is defined as a dispersion of gas in a liquid continuous phase. During foaming, new interfaces are continuously created and populated by surfactants, so, as mentioned above, the adsorption kinetics play an important role. Here it was seen that foamability increased systematically with increasing monoester content, see Figure 19. The monoester-rich sample generated substantially more foam than the 5:1 sample, which in turn produced considerably more foam than the 2:1 sample. The higher foamability is potentially related to the faster adsorption kinetics of the monoesters, enabling more rapid stabilization of newly formed air-water interfaces during foam generation. In addition, the presence of larger and more structurally diverse multiester species may hinder efficient interfacial packing, which could explain the difference in foaming ability also between the 5:1 and 2:1 samples, despite them having similar adsorption kinetics. Together, these effects illustrate how heterogeneity can have a large impact on application-relevant properties despite relatively small differences in equilibrium surface tension.

The results clearly illustrate that heterogeneity in the degree of esterification affects surfactant behavior and particularly the dynamic properties such as adsorption kinetics. From an application perspective, it has large consequence on foamability whereas surface tension at the plateau was comparatively unaffected when equilibrium is reached.

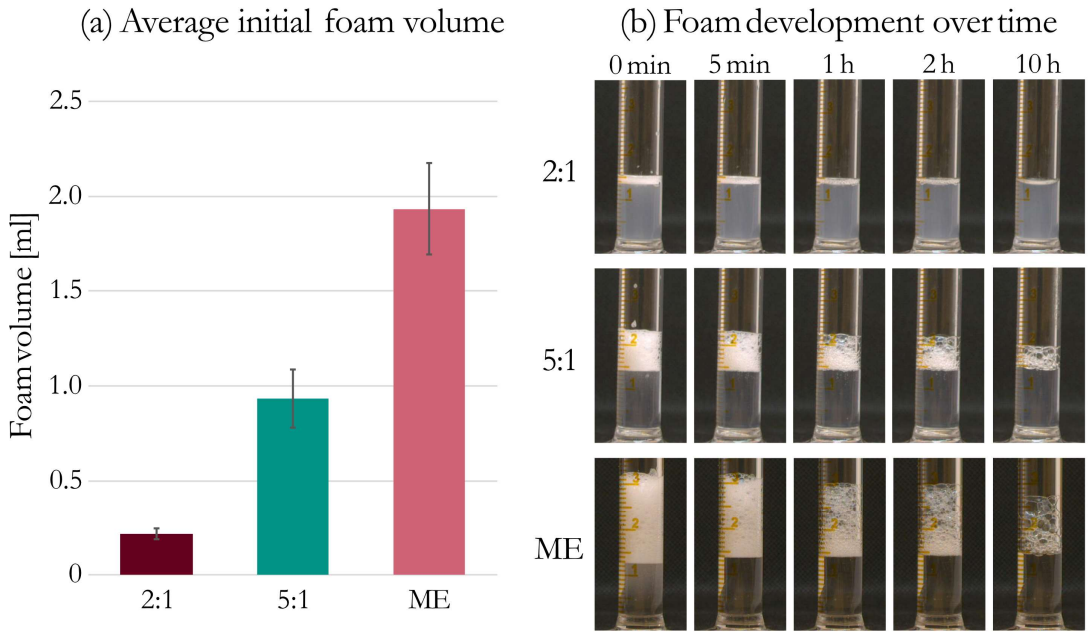

Figure 19. Results from the foam tests. (a) Initial foam volume, average of three replicates with the error bars showing the standard deviation. (b) Foam development over time shown for one representative replicate of each sample. Figure from Paper II.

4.2.2 Structure-property relationship of well-defined model surfactants (Paper III)

To limit the compositional complexity associated with commercial PG esters, a study was conducted where a series of structurally well-defined surfactants were synthesized. Each molecule consisted of a branched triglycerol (G3) headgroup with a single fatty acid chain attached at a defined central position, see Figure 20. The surfactants were prepared from a ketal-protected triglycerol, with only the central hydroxyl available for tail attachment, followed by a deprotection step. Three surfactants with varying tail lengths were evaluated: G3C10, G3C12 and G3C14. This model system made it possible to study the influence of molecular architecture while avoiding the heterogeneity present in conventional PG esters.

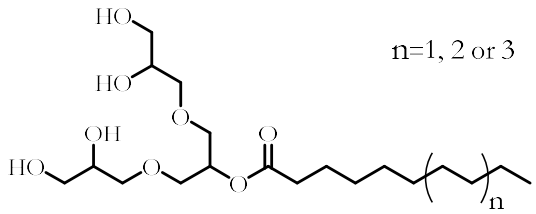

Figure 20. Molecular structure of the branched triglycerol ester surfactants, with either 10, 12 or 14 carbons in the fatty acid tail. Figure from Paper III.

4.2.2.1 Unexpected aggregation behavior

The solution properties of the surfactants revealed behaviors that deviated from classical surfactant expectations. Aqueous solutions of 1 wt % and 20 mM were observed visually, and the solutions of G3C12 and G3C14 formed transparent aqueous solutions with aggregate sizes

consistent with small spherical micelles, as determined from DLS, see Figure 21. In contrast, the shortest-tail surfactant, G3C10, produced slightly cloudy solutions and formed aggregates with diameters approaching 1 µm according to DLS measurements, see Figure 21. This observation is counterintuitive. Conventional arguments related to the CPP would suggest that longer hydrophobic tails should promote the formation of larger self-assembled structures, due to a larger tail volume, and subsequently a larger CPP. Instead, the opposite trend was observed.

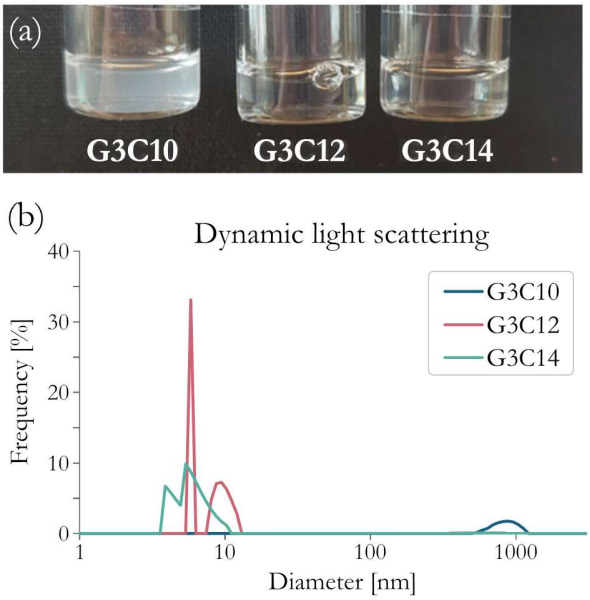

Figure 21. Visual appearance of 20 mM aqueous solutions (a), and volume weighed aggregate size distributions determined with DLS at 1.5 x CMC (b). Both results show that the G3C12 and G3C14 surfactants form transparent, micellar solutions in water, whereas the G3C10 surfactant gives an opalescent appearance with larger aggregates. Figure from Paper III.

A comparison can be done between G3C10 and the ME sample described in section 4.2.1, which also consists of monoesters of mainly triglycerol with 10 carbons in the tail. In contrast to G3C10, the ME sample appeared completely transparent in 1 wt% aqueous solution. The main difference between the two surfactant samples is the heterogeneity of the head group in ME, as it is a mixture both in terms of number of glycerol units, the way these glycerol units are attached to each other, and the tail position on the head. The G3C10 surfactant, on the other hand, has a well-defined branched head-group structure. Whether it is the branched structure of the G3C10 head group, the polydispersity in the ME sample or something different that is the cause of the difference in solution behavior has however not been established, and more work would be needed to understand the underlying cause.

4.2.2.2 Adsorption at the air-water interface

Surface tension measurements showed that the CMC followed the expected trend, decreasing with increasing tail length according to equation 1, see Figure 22 and Table 6. This behavior is consistent with classical surfactant theory. The surface tension at the plateau, however, showed an opposite trend compared to what is typically observed. Rather than decreasing with increasing tail length, the plateau surface tension increased, and the G3C10 surfactant exhibited the lowest surface tension despite having the shortest hydrophobic chain. This was unexpected, and should be investigated further.

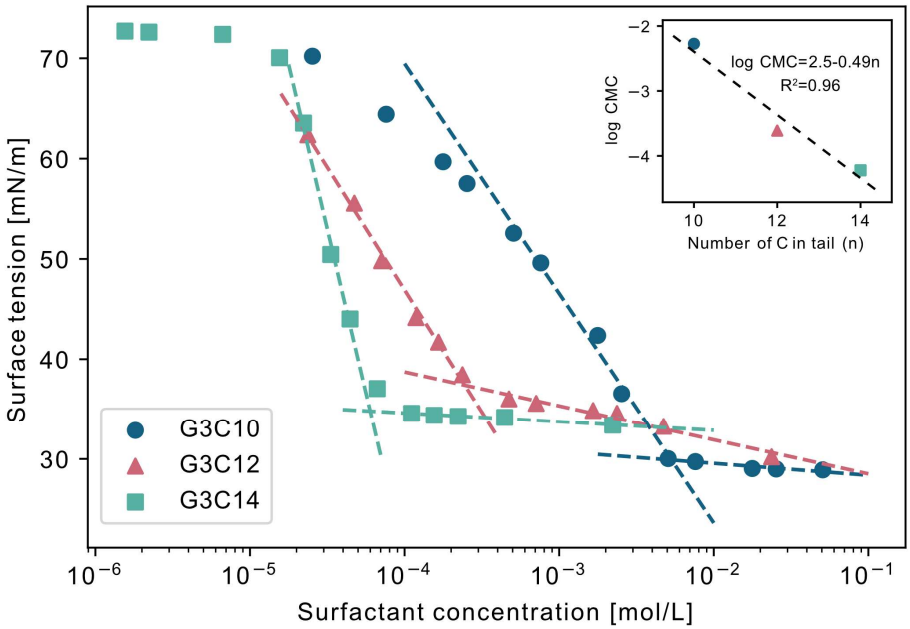

Figure 22. Surface tension isotherms of the three surfactants. The CMC follows the expected trend according to Equation 1, as can be seen in the insert, where log CMC vs number of carbons in the tail (n) is shown. Figure from Paper III.

Table 6. Surface tension at the plateau and critical micelle concentration (CMC) of the three branched triglycerol surfactants.

	Surface tension [mN/m]	CMC [mM]
G3C10	29±1	5.3
G3C12	30±1	0.25
G3C14	33±1	0.060

Again, a comparison can be done between G3C10 and the ME sample. The CMC of G3C10 is 0.21 wt%, which is considerably higher than that estimated for ME (0.03 wt%). Notably, the ME sample was made from a polydisperse PG sample, consisting mainly of di-, tri- and tetraglycerol, and one hypothesis is that the polydispersity of the ME head groups, with a considerable number of diglycerol, gives a fraction of the molecules with a smaller hydrophilic head, which might lower the CMC of the mixture, as head group size has a moderate effect on CMC [9]. It should however be noted that the CMC values were taken after different equilibration times, and might therefore not be directly comparable. The surface tension at the plateau of ME was 26 mN/m, compared to 29 mN/m for G3C10. One hypothesis is that the branched structure of the G3C10 head group is bulkier than those of ME, and cannot pack as efficiently at the interface. It has also been seen before, both in ethoxylated nonionic surfactants [28] and other PG surfactants [29], that a wider distribution of the head group size has been beneficial for surface tension reduction, as was discussed in section 2.1.2.3.

4.2.2.3 Adsorption at solid surfaces

Further evidence for the atypical self-assembly behavior was observed when studying the adsorption at solid surfaces using QCM-D, see Figure 23. The longest-tail surfactant, G3C14, displayed adsorption characteristics consistent with conventional surfactant behavior, with adsorbed masses of 200 ng/cm^2 and 270 ng/cm^2 on hydrophobic and hydrophilic surfaces, respectively. This corresponds approximately to values usually measured for a monolayer and a bilayer with interdigitated tails, respectively [78-80].

In contrast, G3C10 exhibited much larger adsorbed masses, of ca 600-700 ng/cm^2 and 4000 ng/cm^2 on hydrophobic and hydrophilic surfaces, respectively. The shift in dissipation was also substantially higher than expected. These observations indicate the formation of large, soft, adsorbed structures rather than simple surfactant mono- or bilayers. The behavior is in agreement with the formation of larger aggregates observed during the study of the solution behavior.

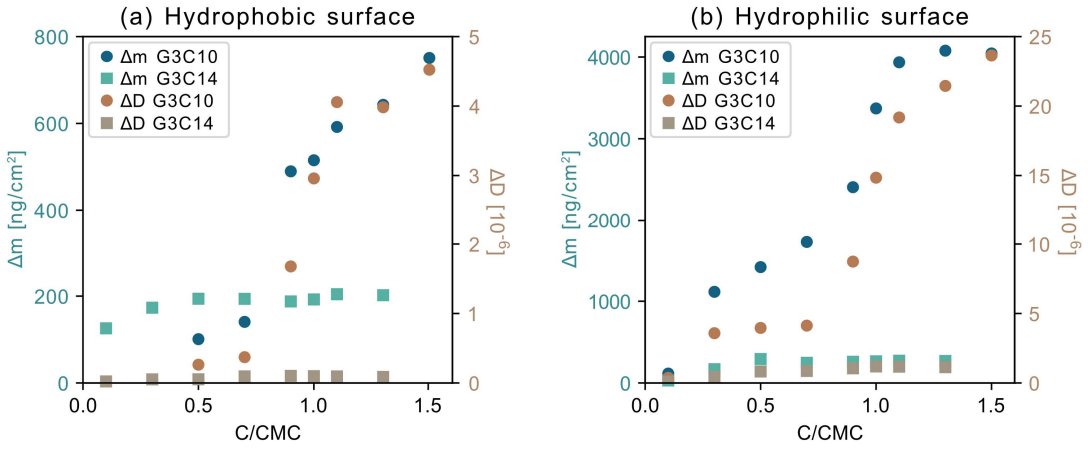

Figure 23. Results from QCM-D measurements, with the adsorbed mass (Δm) and dissipation change (ΔD) on a (a) hydrophobic surface and (b) hydrophilic surface. Figure from Paper III.

4.2.2.4 Role of headgroup interactions

Taken together, the results suggest that even though the CMC follows the expected trend when increasing surfactant tail length, the surfactant with the shortest tail seems to self-assemble into larger aggregates, which is unexpected. One hypothesis to explain this behavior is that extensive intermolecular hydrogen bonding between the head groups reduces the effective area per headgroup, and that this interaction becomes predominant when the tail length is short. Such a reduction could increase the CPP and could explain the tendency of G3C10 to form large self-assembled structures. The effect appears strongest for the shortest-tail surfactant, where the hydrophobic contribution is least able to counterbalance these headgroup interactions. A complete understanding of these observations requires consideration not only of the hydrophobic contribution of the surfactant, but also of secondary interactions that arise once the surfactants are in an aggregated state, as these interactions can influence both the final aggregate structure and the overall behavior of the surfactant in solution. There are, however, some unidentified impurities remaining in the three surfactants. Therefore, one cannot unambiguously state whether or not the observed turbidity is a result of the surfactant structure alone.

4.2.3 Structure-property relationships in polyglycerol ester surfactants

The two studies presented in this chapter provide complementary perspectives on the role of molecular structure in polyglycerol ester surfactants. The first demonstrates that heterogeneity in degree of esterification significantly influences solution behavior, adsorption kinetics and foamability. The second shows that even when heterogeneity is removed, the underlying molecular architecture can lead to unexpected behavior that is not captured by conventional surfactant models.

Together, these findings highlight both the importance and the challenge of establishing quantitative structure-property relationships for PG ester surfactants. The inherent heterogeneity of commercial materials affects functional performance directly, while the behavior of well-defined model compounds reveals molecular interactions that may be masked in heterogeneous systems. Consequently, both approaches are useful: characterization of heterogeneity to understand real industrial surfactants, whereas model systems can be used to identify the fundamental molecular mechanisms that govern their behavior.

4.3 Machine learning-based analysis of heterogeneity in colloidal gold (Paper IV)

The previous chapters focused on identifying heterogeneity in molecular systems and understanding how heterogeneity influences surface properties of surfactants. As the complexity of a material increases, however, its characterization becomes increasingly challenging. One potential approach to address this challenge is the use of machine learning (ML). ML enables the extraction of complex patterns directly from large datasets, without requiring explicit analytical relationships between measured characterization data and the underlying structural parameters.

To explore the potential of such approaches for the characterization of heterogeneous materials, convolutional neural networks (CNNs) were evaluated as a tool for determination of the structural heterogeneity of gold nanoparticles (AuNPs) from UV-vis spectra. CNNs were expected to be well suited for this purpose, as they have proven effective in identifying patterns in complex datasets.

Gold nanoparticles and their optical properties were chosen as a model system due to the well-established relationship between particle structure and spectroscopic response. However, the motivation of this work extends beyond nanoparticle characterization, and serves as an example of how ML can be used to extract structural information about complex and heterogeneous materials from spectroscopic data. In the future, similar approaches could potentially be applied to other heterogeneous systems, including surfactants and their building blocks, as discussed later in this thesis.

4.3.1 Prediction of size distributions from noise-free spectra

For heterogeneous materials, the objective is often not simply to determine an average property but rather to characterize an entire distribution. In the case of AuNPs, structural heterogeneity can be described by particle size distributions. For spherical particles the diameter distribution is the main descriptor, and for rod-shaped particles both the distributions of the rod diameter and aspect ratio (AR), where the AR is the ratio between the length and diameter should be described. As was discussed previously in this thesis, these distributions determine the optical properties of the material but are difficult to extract directly from UV-vis spectra.

Large datasets of simulated UV-vis spectra were generated for spherical and rod-shaped AuNPs with varying size distributions and noise levels, according to the methodology described in

section 3.8. 1D CNN models were then trained to predict the underlying distribution parameters directly from the spectra. Rather than predicting only an average particle size, the models were designed to determine both the mean and the standard deviation (as a measure of polydispersity) of the distributions, thereby providing information about heterogeneity.

For the data without added noise, the trained CNN models were able to accurately predict the parameters that define the AuNP distributions. For spherical particles, both the average diameter and the standard deviation of the particle size distribution could be predicted with high accuracy. The coefficient of determination (R^2) was 0.999 for the average diameter and 0.976 for the diameter polydispersity, see Figure 24.

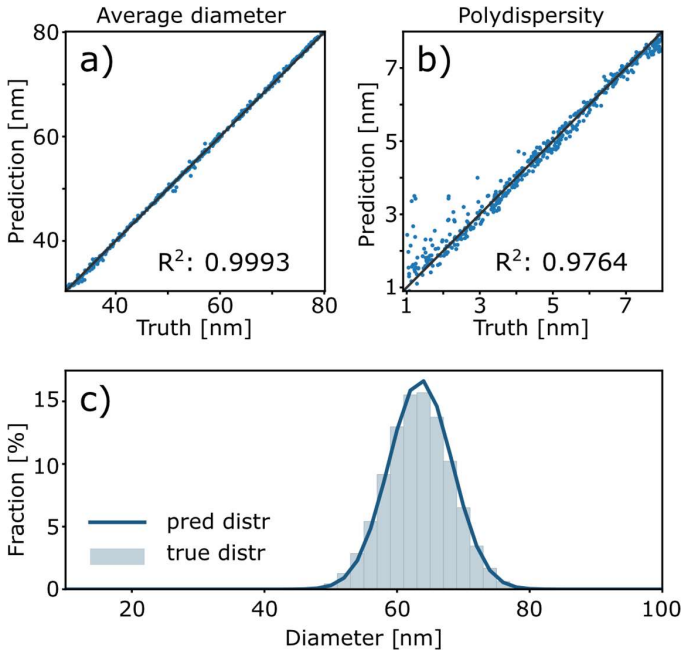

Figure 24. Predictions compared to true size distribution parameters in the test set for the spherical AuNPs, with average diameter in (a), and the standard deviation of the distribution in (b) as a measure of the polydispersity. In (c), one representative example is shown with the predicted and true size distributions. Figure from Paper IV [83].

The same approach was extended to rod-shaped particles, which represent a more complex system in terms of heterogeneity, because two independent distributions must be considered simultaneously: the diameter distribution and the aspect ratio distribution. Despite this increased complexity, the CNN successfully predicted all four distribution parameters. Particularly accurate predictions were obtained for the average diameter and aspect ratio as well as the aspect ratio polydispersity, all with coefficients of determination above 0.99, see Figure 25.

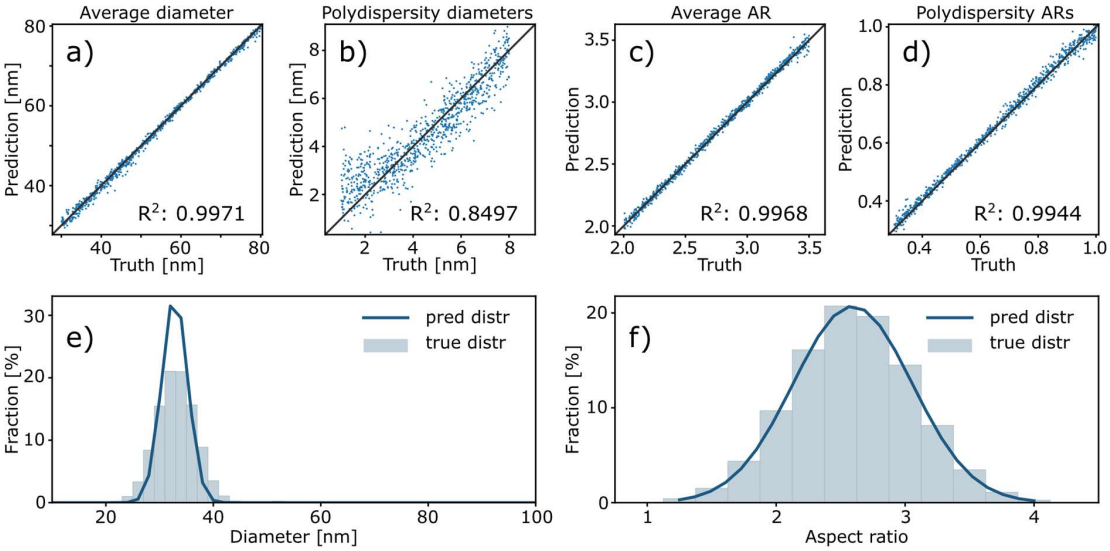

Figure 25. Predictions compared to true size distribution parameters in the test set for the rod-shaped AuNPs, with average diameter in (a), standard deviation of the diameter distribution in (b), average AR in (c) and standard deviation of the AR distribution in (d). In (e) and (f), one representative example is shown with the predicted and true distributions of diameter and AR, respectively. Figure from Paper IV [83].

These results demonstrate that machine learning can extract detailed information about AuNP sizes from UV-vis spectra, and that CNNs represent a solid approach. However, these predictions were done on simulated data that are perfectly smooth, which is typically not the case for experimental data. Therefore, as a tougher challenge for the models, datasets with varying amounts of added noise were used as the next step, as these datasets can be regarded as more representative of real-world systems.

4.3.2 Predictions on data with added noise

To the simulated spectra described in the previous section, normal-distributed noise was added, with varying standard deviations in the range 0-0.03, where a high standard deviation gives a more intense noise. Illustrations of different noise levels are shown at the bottom of Figure 26 and Figure 27. While the CNNs could predict all size distribution parameters accurately for noise-free spectra, the prediction accuracy of diameter polydispersity rapidly deteriorated when noise was introduced, especially for rod-shaped particles, as shown in Figure 26 and Figure 27. Variations in diameter polydispersity produce only subtle changes in the UV-vis spectra, and this is likely making these features increasingly difficult to distinguish as the noise level increases. In contrast, the other parameters, which have a stronger influence on the spectral response, could still be predicted with high accuracy even at the highest investigated noise levels.

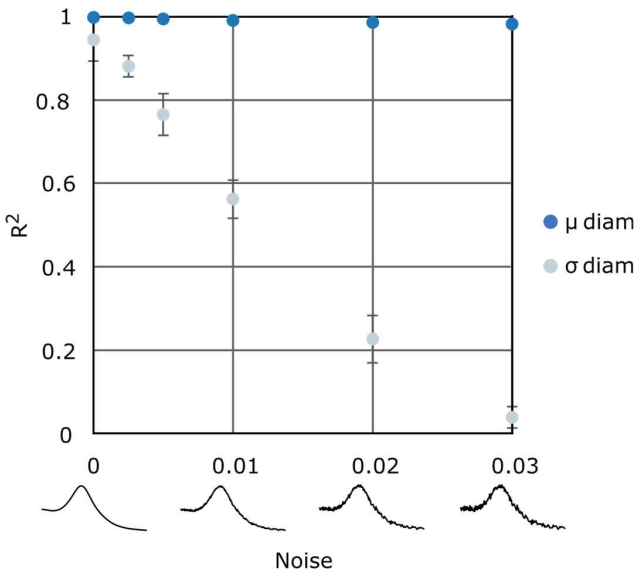

Figure 26. Effect of increasing noise on model performance for the spherical AuNPs, quantified by the coefficient of determination (R^2). Results are shown as a function of the standard deviation of the added normal distributed noise, with example spectra of different noise levels shown at the bottom. The parameters μ and σ denote the mean and standard deviation of the size distribution, respectively. Error bars show the standard deviation of three replicates. Figure from Paper IV [83].

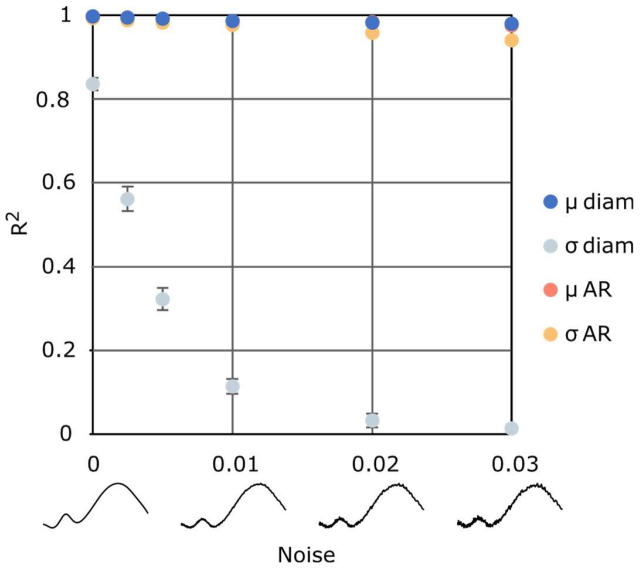

Figure 27. Effect of increasing noise on model performance for the rod-shaped AuNPs, quantified by the coefficient of determination (R^2). Results are shown as a function of the standard deviation of the added normal distributed noise, with example spectra of different noise levels shown at the bottom. The parameters μ and σ denote the mean and standard deviation of the size distribution, respectively. Error bars show the standard deviation of three replicates. Figure from Paper IV [83].

These results highlight an important consideration when using ML to characterize heterogeneous materials: the ability to quantify heterogeneity depends not only on the performance of the model itself, but also on how strongly the underlying heterogeneity is reflected in the data. Structural variations that produce only minor changes in the experimental response may be inherently challenging to resolve, particularly in the presence of noise. Nevertheless, the CNN models demonstrated robust performance for most of the investigated size distribution parameters, indicating that machine learning is a promising approach for extracting quantitative information about heterogeneity from spectroscopic data.

4.3.3 Toward machine learning-assisted characterization of other complex materials

This work demonstrates that machine learning can be used as a powerful tool for characterizing heterogeneity. Rather than relying on explicit physical models to fit experimental data, the CNN learned relationships between spectral features and size distributions directly from the data. Although demonstrated here for gold nanoparticles, the broader significance lies in the methodology. As has been discussed previously in this thesis, many materials are mixtures rather than single molecules. In the previous chapters of this thesis, polyglycerol and polyglycerol ester surfactants were shown to consist of distributions in degree of polymerization, connectivity, branching and degree of esterification. It was shown in **Paper I** that there is a clear correlation between the constitution of PG and its NMR spectrum, although challenging to determine. ML could be a useful tool in extracting this information in the future. One challenge, however, lies in the large amount of data needed to train a model for this purpose. For a supervised deep learning algorithm such as the CNN model used in **Paper IV**, a large number of spectra with *known* composition would be needed for the training. A potential strategy to tackle this challenge could be to use pure oligomeric PG standards, like the ones synthesized by Cassel et al. [92], and make mixtures of known composition in terms of the glycerol substructures. The NMR spectra of these mixtures, together with the know composition, would then form the input-output data pairs needed to train a ML model.

Another potential use of ML for heterogeneous materials could be to bypass the need for structural characterization. Instead of using spectra for determination of the structure of a surfactant, and then comparing the structure to its properties, one could potentially go directly from the spectra to the properties by using ML. For example, one could train a supervised ML model on data pairs consisting of a spectrum as input (for example NMR or FTIR spectra), and the value of some desired property as the output (for example the CMC, surface tension at the plateau or foam height). If the structural features governing the property of interest are encoded in the measured spectrum, it should in principle be possible to predict that property directly from the spectral data. Such an approach could bypass the need for detailed, complex, and time-consuming characterization of the underlying molecular heterogeneity, while still capturing the structural information most relevant to the macroscopic behavior of the system.

5 Conclusions

The purpose of this thesis was to deepen the understanding of heterogeneity in surface chemistry, and to investigate strategies for characterization of heterogeneous systems and their surface properties. Particular focus was placed on polyglycerol ester surfactants, where heterogeneity exists at multiple structural levels, ranging from the constitution of the polyglycerol building block to the composition of the final surfactant product in terms of degree of esterification. In parallel, machine learning was explored as a data-driven approach for characterization of heterogeneity in colloidal gold nanoparticle systems, and extraction of complex features inherently encoded in spectral data. The studies demonstrate that heterogeneity can strongly influence material properties and that detailed characterization is valuable for establishing meaningful structure-property relationships.

In **Paper I**, a methodology was developed for characterization of the constitutional heterogeneity of polyglycerol materials using quantitative and multidimensional NMR spectroscopy. The work showed that commercial polyglycerols cannot be adequately described solely by their average degree of polymerization. Instead, they consist of a mixture of molecules containing structurally distinct glycerol subunits connected through different linkage patterns. By extending previous classification approaches a more detailed description of the molecular architecture could be achieved. The new classification, in combination with the quantification methodology, revealed that the complexity of the material increases with increasing oligomer size and demonstrated that quantitative ^{13}C NMR provides valuable information regarding the heterogeneity of these surfactant building blocks. Furthermore, a simplified approach based on quantitative ^{1}H NMR was shown to provide a practical alternative for routine characterization when detailed substructural information is not required. This NMR methodology could be a valuable tool for determining structure-property relationships.

In **Paper II**, the influence of molecular heterogeneity on surfactant properties was investigated by focusing on the presence of di- and higher esters of polyglycerol. The results showed that heterogeneity in degree of esterification significantly affects solution behavior, adsorption kinetics, and foaming performance. The sample enriched in monoesters adsorbed more rapidly at the air-water interface and generated substantially more foam than more heterogeneous mixtures, while equilibrium surface tension was comparatively less affected. These findings demonstrate that compositional heterogeneity can strongly influence dynamic and application-relevant properties even when equilibrium properties remain similar.

In **Paper III**, structurally well-defined branched triglycerol ester surfactants were used to address the challenge of establishing structure-property relationships in highly heterogeneous surfactant systems. By minimizing compositional heterogeneity and controlling both the headgroup architecture and tail position and length, it became possible to isolate the effects of specific molecular features on surfactant behavior. The results showed that while the CMC followed the expected dependence on tail length, several other properties did not conform to conventional surfactant trends. In particular, the shortest-tail surfactant, i.e. the most hydrophilic, exhibited unexpectedly large aggregates in aqueous solution, unusually low equilibrium surface tension, and extensive adsorption at solid surfaces, compared to the homologues with longer tails. One hypothesis is that intermolecular hydrogen bonding between the branched head groups plays an important role in the self-assembly behavior and contributes to the deviation from conventional surfactant trends. These findings highlight the importance of studying well-defined model compounds alongside heterogeneous commercial mixtures to reveal the fundamental molecular interactions governing surfactant behavior.

In **Paper IV**, heterogeneity was examined from a different perspective, focused on colloidal gold nanoparticle systems. Convolutional neural network models were successfully trained to predict particle size distributions directly from UV-vis spectra. The models accurately determined both average size and polydispersity, demonstrating that machine learning can extract information about heterogeneity that is difficult to obtain using conventional spectral analysis. Although prediction accuracy decreased for parameters whose spectral signatures were obscured by factors such as noise, the overall results demonstrate that machine learning is a powerful tool for characterizing complex distributions in heterogeneous materials.

Taken together, the work presented in this thesis demonstrates that heterogeneity can be analyzed using either data-driven approaches or combined quantitative and multidimensional NMR methods. Furthermore, this work underscores that heterogeneity can play an important role in the surface properties and performance of polyglycerol ester surfactants, and that well-defined model systems can be used to identify the underlying mechanisms by which molecular structure affects material properties.

Several opportunities remain for further investigation. An important next step is to extend the NMR methodology developed for polyglycerol characterization to esterified products, including the identification of tail positions. The unexpected behavior observed for the branched triglycerol ester surfactants and the role of headgroup interactions in determining aggregate structure would benefit from further studies. Finally, the machine-learning approach demonstrated for gold nanoparticles could potentially be extended to heterogeneous surfactant systems, enabling prediction of structural distributions or material properties directly from spectroscopic data.

Acknowledgements

This work was funded by the Swedish Research Council (Vetenskapsrådet) and Vinnova.

First of all, I would like to thank my supervisors: Romain Bordes, for your optimism and for helping me find my way back on track whenever I lost myself in the details. My co-supervisor, Lars Evenäs, for taking the time to explain, for your attention to detail, and for teaching me the importance of positive feedback. Thank you both for helping me grow as a scientist.

I would also like to thank my examiner, Per-Anders Carlsson, for your advice regarding my PhD journey and for your objective perspective and encouragement.

Big thanks to Krister Holmberg, for your interest in my work, for input regarding my papers and for interesting discussions.

Thanks also to all my co-authors of the papers included in this thesis. I would like to thank Hannes for nice discussions and for his help with the purification of my surfactants.

To all my former and current colleagues at Applied Chemistry, thank you for making this such a pleasant place to work. I have appreciated the many lunches and coffee breaks, and the friendly, helpful atmosphere throughout the division. Special thanks to Maja, Smaragda, Moa, Josmary, Andreas, Hugo and Nasim, for your friendship and support.

A big thanks to all my students, both those I have supervised in BSc and MSc thesis projects and those I have taught in courses, whose enthusiasm made teaching very enjoyable. Special thanks to Moa, Eric, Nidhi and Molly, whose work contributed with valuable results to this thesis.

I want to thank my friends and family for their support throughout these years. A special thanks to my parents for always being there for me. Thanks to my biggest cheerleaders, farmor och farfar, I know you would have been proud.

To my son Hugo, for being such a joy in my life. And most of all, a heartfelt thanks to my husband Fredrik, for your steady support and love.

Frida Bilén
Mölndal, August 2026

References

1. Britannica Academic, *Heterogeneous reaction*, in *Encyclopædia Britannica*. 1998.
2. Zhang, M. and G. Zhang, *Recent Advances in the Properties and Applications of Polyglycerol Fatty Acid Esters.* Polymers, 2025. **17**(7): p. 879.
3. *Polyglycerol Market Size and Share Industry report, 2033*. [Internet] 2025 [cited 2026 June 03]; Available from: https://www.grandviewresearch.com/industry-analysis/polyglycerol-market.
4. Norn, V., *Polyglycerol Esters*, in *Emulsifiers in Food Technology*. 2014. p. 181-208.
5. Doak, J., et al., *Effect of particle size distributions on absorbance spectra of gold nanoparticles.* Physica E: Low-dimensional Systems and Nanostructures, 2010. **42**(5): p. 1605-1609.
6. Kronberg, B., K. Holmberg, and B. Lindman, *Chapter 1 Types of Surfactants, their Synthesis, and Applications*, in *Surface Chemistry of Surfactants and Polymers*. 2014. p. 1-47.
7. Kronberg, B., K. Holmberg, and B. Lindman, *Chapter 12 Surface and Interfacial Tension*, in *Surface Chemistry of Surfactants and Polymers*. 2014. p. 231-249.
8. Davies, J. *A quantitative kinetic theory of emulsion type, I. Physical chemistry of the emulsifying agent*. in *Gas/Liquid and Liquid/Liquid Interface. Proceedings of the International Congress of Surface Activity*. 1957. Citeseer.
9. Kronberg, B., K. Holmberg, and B. Lindman, *Chapter 4 Surfactant Self-Assembly: General aspects and spherical micelles*, in *Surface Chemistry of Surfactants and Polymers*. 2014. p. 75-94.
10. Klevens, H.B., *Structure and aggregation in dilute solution of surface active agents.* Journal of the American Oil Chemists' Society, 1953. **30**(2): p. 74-80.
11. Rosen, M.J., *Surfactants and interfacial phenomena*. 3 ed. 2004, Hoboken, New Jersey: John Wiley & Sons, Inc.
12. Israelachvili, J.N., D.J. Mitchell, and B.W. Ninham, *Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers.* Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics, 1976. **72**: p. 1525-1568.
13. Kronberg, B., K. Holmberg, and B. Lindman, *Chapter 6 Surfactant Self-Assembly: Beyond the spherical micelle*, in *Surface Chemistry of Surfactants and Polymers*. 2014. p. 113-136.
14. Kronberg, B., K. Holmberg, and B. Lindman, *Chapter 7 Surfactants and Polymers Containing Oxyethylene Groups Show a Complex Behavior*, in *Surface Chemistry of Surfactants and Polymers*. 2014. p. 137-152.
15. Lewis, J.J., *Ch 8 - Polyol Ester Surfactants*, in *Nonionic surfactants - Organic chemistry*, N.M. van Os, Editor. 1998, Marcel Dekker, Inc.: New York.
16. Alberio, C. and L.A.N. Aguirrezábal, *Meta-analysis unravels common responses of seed oil fatty acids to temperature for a wide set of genotypes of different plant species.* Frontiers in Plant Science, 2024. **15**.
17. Ghaffari, M., et al., *Drought-stress induced changes of fatty acid composition affecting sunflower grain yield and oil quality*. Food Science & Nutrition, 2023. **11**(12): p. 7718-7731.
18. Kosswig, K., *Ch 4 - Polyoxyethylene Esters of Fatty Acids*, in *Nonionic surfactants - Organic chemistry*, N.M. van Os, Editor. 1998, Marcel Dekker, Inc.: New York.
19. Edwards, C.L., *Ch 1 - Distribution of the Polyoxyethylene Chain*, in *Nonionic surfactants - Organic chemistry*, N.M. van Os, Editor. 1998, Marcel Dekker, Inc.: New York.
20. von Rybinski, W. and K. Hill, *Ch 2 - Alkyl Polyglycosides*, in *Novel Surfactants - Preparation, Applications, and Biodegradability. Second Edition, Revised and Expanded*, K. Holmberg, Editor. 2003, Marcel Dekker, Inc.: New York.
21. Borisov, O.V., et al., *Toward Understanding Molecular Heterogeneity of Polysorbates by Application of Liquid Chromatography–Mass Spectrometry with Computer-Aided Data Analysis.* Analytical Chemistry, 2011. **83**(10): p. 3934-3942.

22. Frison-Norrie, S. and P. Sporns, *Investigating the Molecular Heterogeneity of Polysorbate Emulsifiers by MALDI-TOF MS*. Journal of Agricultural and Food Chemistry, 2001. **49**(7): p. 3335-3340.
23. Vu Dang, H., et al., *Composition analysis of two batches of polysorbate 60 using MS and NMR techniques*. Journal of Pharmaceutical and Biomedical Analysis, 2006. **40**(5): p. 1155-1165.
24. Chong, C.C., et al., *A review over the role of catalysts for selective short-chain polyglycerol production from biodiesel derived waste glycerol*. Environmental Technology & Innovation, 2020. **19**: p. 100859.
25. Márquez-Alvarez, C., E. Sastre, and J. Pérez-Pariente, *Solid Catalysts for the Synthesis of Fatty Esters of Glycerol, Polyglycerols and Sorbitol from Renewable Resources*. Topics in Catalysis, 2004. **27**(1): p. 105-117.
26. Ciriminna, R., et al., *Glycerol-Derived Renewable Polyglycerols: A Class of Versatile Chemicals of Wide Potential Application*. Organic Process Research & Development, 2015. **19**(7): p. 748-754.
27. Salehpour, S. and M.A. Dubé, *Towards the Sustainable Production of Higher-Molecular-Weight Polyglycerol*. Macromolecular Chemistry and Physics, 2011. **212**(12): p. 1284-1293.
28. Folmer, B.M. and K. Holmberg, *The cross-sectional headgroup area of nonionic surfactants; the influence of polydispersity*. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 2001. **180**(1): p. 187-191.
29. Ishitobi, M. and H. Kunieda, *Effect of chain length distribution on the phase behavior of polyglycerol fatty acid ester in water*. Colloid and Polymer Science, 2000. **278**(9): p. 899-904.
30. Braun, A.C., et al., *Predicting critical micelle concentration and micelle molecular weight of polysorbate 80 using compendial methods*. European Journal of Pharmaceutics and Biopharmaceutics, 2015. **94**: p. 559-568.
31. Huang, X., S. Neretina, and M.A. El-Sayed, *Gold Nanorods: From Synthesis and Properties to Biological and Biomedical Applications*. Advanced Materials, 2009. **21**(48): p. 4880-4910.
32. Dreaden, E.C., et al., *The golden age: gold nanoparticles for biomedicine*. Chemical Society Reviews, 2012. **41**(7): p. 2740-2779.
33. Dreaden, E.C., et al., *Beating cancer in multiple ways using nanogold*. Chemical Society Reviews, 2011. **40**(7): p. 3391-3404.
34. Uusitalo, M., et al., *Gold Nanorods as Photothermal Antibacterial Materials*. ACS Applied Nano Materials, 2025. **8**(7): p. 3661-3670.
35. Lidström Uusitalo, M., *Solid-Supported Gold Nanorods for Antibacterial Applications*, in *Department of Chemistry and Chemical Engineering*. 2025, Chalmers University of Technology: Gothenburg, Sweden.
36. Daniel, M.-C. and D. Astruc, *Gold Nanoparticles: Assembly, Supramolecular Chemistry, Quantum-Size-Related Properties, and Applications toward Biology, Catalysis, and Nanotechnology*. Chemical Reviews, 2004. **104**(1): p. 293-346.
37. Hamamoto, M. and H. Yagyu, *Particle size distribution and Au concentration dependence of the refractive-index sensitivity of LSPR sensors based on gold nanoparticles*. Journal of Nanoparticle Research, 2023. **25**(8): p. 158.
38. Link, S. and M.A. El-Sayed, *Shape and size dependence of radiative, non-radiative and photothermal properties of gold nanocrystals*. International Reviews in Physical Chemistry, 2000. **19**(3): p. 409-453.
39. Jain, P.K., et al., *Calculated Absorption and Scattering Properties of Gold Nanoparticles of Different Size, Shape, and Composition: Applications in Biological Imaging and Biomedicine*. The Journal of Physical Chemistry B, 2006. **110**(14): p. 7238-7248.
40. Link, S. and M.A. El-Sayed, *Size and Temperature Dependence of the Plasmon Absorption of Colloidal Gold Nanoparticles*. The Journal of Physical Chemistry B, 1999. **103**(21): p. 4212-4217.

41. Chithrani, B.D., A.A. Ghazani, and W.C.W. Chan, *Determining the Size and Shape Dependence of Gold Nanoparticle Uptake into Mammalian Cells.* Nano Letters, 2006. **6**(4): p. 662-668.
42. Lohse, S.E. and C.J. Murphy, *The Quest for Shape Control: A History of Gold Nanorod Synthesis.* Chemistry of Materials, 2013. **25**(8): p. 1250-1261.
43. Haiss, W., et al., *Determination of Size and Concentration of Gold Nanoparticles from UV−Vis Spectra.* Analytical Chemistry, 2007. **79**(11): p. 4215-4221.
44. Wang, X. and Y. Cao, *Characterizations of absorption, scattering, and transmission of typical nanoparticles and their suspensions.* Journal of Industrial and Engineering Chemistry, 2020. **82**: p. 324-332.
45. Iqbal, M., et al., *Preparation of gold nanoparticles and determination of their particles size via different methods.* Materials Research Bulletin, 2016. **79**: p. 97-104.
46. Hu, Z.J., et al., *Fast characterization of gold nanorods ensemble by correlating its structure with optical extinction spectral features.* AIP Advances, 2014. **4**(11).
47. Njoki, P.N., et al., *Size Correlation of Optical and Spectroscopic Properties for Gold Nanoparticles.* The Journal of Physical Chemistry C, 2007. **111**(40): p. 14664-14669.
48. Murphy, K.P., *Machine learning: a probabilistic perspective.* 2012: MIT Press.
49. Goodfellow, I., Y. Bengio, and A. Courville, *Deep Learning.* 2016: MIT Press.
50. Schuetzke, J., N.J. Szymanski, and M. Reischl, *Validating neural networks for spectroscopic classification on a universal synthetic dataset.* npj Computational Materials, 2023. **9**(1): p. 100.
51. Nielsen, M.A., *Neural Networks and Deep Learning*. 2015: Determination Press.
52. LeCun, Y., Y. Bengio, and G. Hinton, *Deep learning.* Nature, 2015. **521**(7553): p. 436-444.
53. Meza Ramirez, C.A., et al., *Applications of machine learning in spectroscopy.* Applied Spectroscopy Reviews, 2021. **56**(8-10): p. 733-763.
54. Keras. *Adam.* [cited 2026 July 22]; Available from: https://keras.io/api/optimizers/adam/.
55. Elliott, R.J., et al., *Using neural networks to deduce polymer molecular weight distributions from linear rheology.* Journal of Rheology, 2025. **69**(6): p. 955-972.
56. Angulo, A., et al., *Machine learning enhanced spectroscopic analysis: towards autonomous chemical mixture characterization for rapid process optimization.* Digital Discovery, 2022. **1**(1): p. 35-44.
57. Shiratori, K., et al., *Machine-Learned Decision Trees for Predicting Gold Nanorod Sizes from Spectra.* The Journal of Physical Chemistry C, 2021. **125**(35): p. 19353-19361.
58. Karlik, B., et al., *A Hybrid Machine Learning Model to Study UV-Vis Spectra of Gold Nanospheres.* Plasmonics, 2021. **16**(1): p. 147-155.
59. He, J., et al., *Plasmonic nanoparticle simulations and inverse design using machine learning.* Nanoscale, 2019. **11**(37): p. 17444-17459.
60. Tan, E.X., et al., *Incorporating plasmonic featurization with machine learning to achieve accurate and bidirectional prediction of nanoparticle size and size distribution.* Nanoscale Horizons, 2022. **7**(6): p. 626-633.
61. Glaubitz, C., et al., *Leveraging Machine Learning for Size and Shape Analysis of Nanoparticles: A Shortcut to Electron Microscopy.* The Journal of Physical Chemistry C, 2024. **128**(1): p. 421-427.
62. Friebolin, H., *Basic One- and Two-Dimensional NMR Spectroscopy*. 5 ed. 2011: Wiley-VCH.
63. Claridge, T.D.W., *Chapter 9 - Diffusion NMR spectroscopy*, in *Tetrahedron Organic Chemistry Series*, T.D.W. Claridge, Editor. 2009, Elsevier. p. 303-334.
64. Röding, M., et al., *The gamma distribution model for pulsed-field gradient NMR studies of molecular-weight distributions of polymers.* Journal of Magnetic Resonance, 2012. **222**: p. 105-111.
65. Berry, J.D., et al., *Measurement of surface and interfacial tension using pendant drop tensiometry.* Journal of Colloid and Interface Science, 2015. **454**: p. 226-237.
66. Ai, S. and M. Ishitobi, *Effects of the number of fatty acid residues on the phase behaviors of decaglycerol fatty acid esters.* Journal of Colloid and Interface Science, 2006. **296**(2): p. 685-689.

67. Sagitani, H., Y. Hayashi, and M. Ochiai, *Solution properties of homogeneous polyglycerol dodecyl ether nonionic surfactants.* Journal of the American Oil Chemists' Society, 1989. **66**(1): p. 146-152.
68. Kunieda, H., et al., *Cloud and HLB Temperatures of Polyglycerol Didodecanoate Solutions.* Journal of Oleo Science, 2002. **51**(6): p. 379-386.
69. Hollas, M.J., *Modern Spectroscopy*. 4 ed. 2004, Chichester, England: John Wiley & Sons Ltd.
70. Schad, T., et al., *Less is more: Unstable foams clean better than stable foams.* Journal of Colloid and Interface Science, 2021. **590**: p. 311-320.
71. Bellman, C., et al., *Dynamic and Electrophoretic Light Scattering - Guidelines for particle-size analysis and zeta-potential determination*. 2022: Anton Paar GmbH.
72. Malvern Instruments Limited, *Dynamic Light Scattering: An Introduction in 30 Minutes*. 2017.
73. Johannsmann, D. and I. Reviakine, *Quartz crystal microbalance with dissipation monitoring for studying soft matter at interfaces.* Nature Reviews Methods Primers, 2024. **4**(1): p. 63.
74. Sauerbrey, G., *Verwendung von Schwingquarzen zur Wägung dünner Schichten und zur Mikrowägung.* Zeitschrift für Physik, 1959. **155**(2): p. 206-222.
75. Höök, F. and B. Kasemo, *The QCM-D Technique for Probing Biomacromolecular Recognition Reactions*, in *Piezoelectric Sensors*, C. Steinem and A. Janshoff, Editors. 2007, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 425-447.
76. Bordes, R. and F. Höök, *Separation of Bulk Effects and Bound Mass during Adsorption of Surfactants Probed by Quartz Crystal Microbalance with Dissipation: Insight into Data Interpretation.* Analytical Chemistry, 2010. **82**(21): p. 9116-9121.
77. Ulman, A., *Formation and Structure of Self-Assembled Monolayers.* Chemical Reviews, 1996. **96**(4): p. 1533-1554.
78. Mivehi, L., R. Bordes, and K. Holmberg, *Adsorption of Cationic Gemini Surfactants at Solid Surfaces Studied by QCM-D and SPR: Effect of the Rigidity of the Spacer.* Langmuir, 2011. **27**(12): p. 7549-7557.
79. Mivehi, L., R. Bordes, and K. Holmberg, *Adsorption of cationic gemini surfactants at solid surfaces studied by QCM-D and SPR—Effect of the presence of hydroxyl groups in the spacer.* Colloids and Surfaces A: Physicochemical and Engineering Aspects, 2013. **419**: p. 21-27.
80. Bordes, R., J. Tropsch, and K. Holmberg, *Adsorption of Dianionic Surfactants Based on Amino Acids at Different Surfaces Studied by QCM-D and SPR.* Langmuir, 2010. **26**(13): p. 10935-10942.
81. Liu, G. and G. Zhang, *Basic Principles of QCM-D*, in *QCM-D Studies on Polymer Behavior at Interfaces*, G. Liu and G. Zhang, Editors. 2013, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 1-8.
82. Chollet, F. *Keras*. 2015 [cited 2023 6-jan]; Available from: https://keras.io.
83. Bilén, F., et al., *Machine Learning-Based Interpretation of Optical Properties of Colloidal Gold with Convolutional Neural Networks.* The Journal of Physical Chemistry C, 2024. **128**(33): p. 13909-13916.
84. Imran ul-haq, M., et al., *Influence of architecture of high molecular weight linear and branched polyglycerols on their biocompatibility and biodistribution.* Biomaterials, 2012. **33**(35): p. 9135-9147.
85. Stiriba, S.-E., H. Kautz, and H. Frey, *Hyperbranched Molecular Nanocapsules: Comparison of the Hyperbranched Architecture with the Perfect Linear Analogue.* Journal of the American Chemical Society, 2002. **124**(33): p. 9698-9699.
86. Schubert, C., et al., *Systematic Variation of the Degree of Branching (DB) of Polyglycerol via Oxyanionic Copolymerization of Glycidol with a Protected Glycidyl Ether and Its Impact on Rheological Properties.* Macromolecular Chemistry and Physics, 2018. **219**(1): p. 1700376.
87. Tryznowski, M., Z. Żołek-Tryznowska, and J. Izdebska-Podsiadły, *The wettability effect of branched polyglycerols used as performance additives for water-based printing inks.* Journal of Coatings Technology and Research, 2018. **15**(3): p. 649-655.

88. Yamazaki, M., et al., *Effect of Branching Degree of Dendritic Polyglycerols on Plasma Protein Adsorption: Relationship between Hydration States and Surface Morphology.* Langmuir, 2021. **37**(28): p. 8534-8543.
89. Sunder, A., et al., *Controlled Synthesis of Hyperbranched Polyglycerols by Ring-Opening Multibranching Polymerization.* Macromolecules, 1999. **32**(13): p. 4240-4246.
90. Spears, B.R., et al., *Aqueous Epoxide Ring-Opening Polymerization (AEROP): Green Synthesis of Polyglycidol with Ultralow Branching.* Macromolecules, 2016. **49**(6): p. 2022-2027.
91. Rokicki, G., et al., *Hyperbranched aliphatic polyethers obtained from environmentally benign monomer: glycerol carbonate.* Green Chemistry, 2005. **7**(7): p. 529-539.
92. Cassel, S., et al., *Original Synthesis of Linear, Branched and Cyclic Oligoglycerol Standards.* European Journal of Organic Chemistry, 2001. **2001**(5): p. 875-896.
93. Spiga Nord S.p.A. *About polyglycerols*. c 2006-2024 [cited 2026 June 16]; Available from: https://www.spiganord.com/spiga-nord-polyglycerols-pure-vegetable-kosher-grade-polyglycerine/#polyglycerols.
94. Stjerndahl, M., *Biodegradable Surfactants Containing Hydrolyzable Bonds*, in *Department of Chemical and Biological Engineering*. 2005, Chalmers University of Technology: Gothenburg, Sweden.
95. Shrestha, L.K., et al., *Aqueous foam stabilized by dispersed surfactant solid and lamellar liquid crystalline phase.* Journal of Colloid and Interface Science, 2006. **301**(1): p. 274-281.