Sampling Error of TOC swab in Pharmaceutical Cleaning Verification

Jul-Jørgensen, Isabella; Hundahl, Christian A; Skibsted, Erik; Gernaey, Krist V

Published in:
Journal of Pharmaceutical and Biomedical Analysis

Link to article, DOI:
10.1016/j.jpba.2022.114763

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Keywords: Total organic carbon swab, Pharmaceutical cleaning verification, Monte Carlo simulation

Abstract

Cleaning verification is a critical process for patient safety in pharmaceutical manufacturing in order to keep cross-contamination below acceptable limits. A common cleaning verification method is the total organic carbon (TOC) swab. Others have studied the variances of different factors on the TOC swab in order to establish the best swab method. This paper attempts to quantify the sampling error of the TOC swab in the actual sampling situation using simulation. The study investigates the variability on the drug product recovery due to different analysts, concentration, steel finish and position, as well as the estimation of the given swab area. The results demonstrate that the sampling error leads to a large variation in TOC results. For areas estimated in the laboratory, it leads to an increase in limit of detection, LOD, with 60%, while for areas estimated in a tank, the LOD cannot be determined due to the large heteroscedasticity. Thus, this paper is also to be considered as an invitation to discuss and further investigate the TOC sampling error in the pharmaceutical industry.

1. Introduction

Cleaning verification is an essential process in pharmaceutical manufacturing, as it ensures that the maximal allowable carry-over (MACO) or cross-contamination limit between batches is not compromised. Cleaning verification is thus critical for patient safety. A common procedure of cleaning verification is the total organic carbon (TOC) swab method, where the surface of the cleaned equipment is swabbed [1]. Afterwards, the swab stick is analysed for TOC using a TOC analyser. While the TOC analysis itself is a reliable method [2], the sampling is flawed. Several studies have sought to highlight different factors that might influence the cleaning verification result, such as different analysts swabbing [3,4], diluent/solvent composition [4-6], swabbing direction/technique [4,6], type of swabs used [4,6], surface finish, swab area, spiked amount [4], location for TOC sample preparation, [5] and the effect of cleanliness of stainless steel surface on sample recovery [7]. However, these studies mainly focus on how to establish the best swabbing procedure to be used. Sources of variation remain in the actual sampling that will lead to errors, no matter how well-optimised the procedure is beforehand. This is shown in Fig. 1 in the Supplemental material. Unfortunately, this sampling error is often ignored, and validation of the TOC swab method only relies on the analytical error, whereas it should be validated with regards to the total error, TE, given by TE = SE + AE, where SE is sampling error and AE is analytical error [8]. The Parental Drug Association (PDA) highlights that it is essential for cleaning validation/verification that appropriate sampling methods are utilised [1]. Understanding and controlling sampling bias must therefore be a driving force in pharmaceutical cleaning. This study highlights the sampling error that still exists, even when using an established method after optimisation of the swabbing procedure in the laboratory. To the authors’ knowledge, this study is also the first that uses Monte Carlo simulation to give a quantitative estimate of the variability in TOC results in a real sampling situation due to the influence of multiple factors at once. The factors of interest in this study are factors that vary in a real sampling situation and thus influence the result. These factors are variations due to 1) analyst variability on both recovery and area 2) swabbing position, 3) surface finish and 4) drug product concentration. The analyst variability is a substantial factor for the recovery, which is also noted in other studies [3,4]. However, this study also examines how different persons swab different sized areas, which leads to an error in TOC results. Furthermore, the influence of the swabbing position (i.e., top, bottom or side of tank) on the drug product recovery and area estimation was investigated. In a pharmaceutical production the surface finish of the stainless-steel tanks might vary...
depending on the time since the last electropolishing. Though one study investigated the effect of surface finish [4], the lowest value was well below what is typically used for a pharma grade stainless steel, e.g., 316S Stainless Steel [1,9]. Thus, in this study surface finish is investigated in the range of 0.1–1 RA µm, (equivalent to finish #4–#8 in USA [10]), to comply with the GMP requirement that the surface should be easy to clean [11]. Furthermore, influence of the drug product concentration on recovery was also investigated, as the recovery factor is determined at a given drug product concentration. Though some studies have investigated this effect as well [4,6], the concentration range considered was only 90–110% and 50–200% of the residual acceptance limit (RAL). This limit is based on the MACO [12]. The establishment of this limit is not discussed here, but the reader is referred to [13] for more information. Preferably, recovery should be determined at the limit of detection (LOD) or limit of quantification (LOQ) [13]. This study covers a range of 50–500% of the RAL, as this encompasses the range from the current LOD up to the often-used visibly clean limit [14]. A simulation approach was chosen because no method currently exists to non-destructively measure the concentration and to calculate the actual area swabbed simultaneously. The simulation is therefore a combination of different experiments in the laboratory combined with image analysis of areas swabbed with blue dye both in a laboratory setting and a formerly used production tank.

2. Materials and Methods

2.1. Materials

The drug product used was an insulin solution drug product. The swab sticks used were Texwipe TX 761, polyester swab. The TOC vials were clear borosilicate vials with < 10 ppb TOC from C&G containers. Purified water was used as the solvent in the TOC vials, along with phosphoric acid 85%. Blue dye was used for the area estimations.

2.2. Instrument and Software

The TOC samples were analysed with a M9 laboratory analyser from Sievers. The area estimations were processed to numerical values with Python version 3.9.4 using tkinter and PIL packages with a pixel threshold of 150. The script applies a threshold to make a new image, where each pixel is either black or white. Whether a given grey pixel becomes black or white is determined by this threshold. The script then counts the total number of black pixels. The digital scans are all taken in the exact size of a standard A4 sheet, allowing direct conversion between the number of pixels and cm². While the choice of the threshold influences the absolute value of cm² swabbed, this uncertainty is fairly small compared to the variation in swabbed areas. The script was validated by reference scans with known areas of 100 cm². The actual data processing and simulation were performed with Matlab R2021a [15].

2.3. Method

The recovery results were obtained by performing TOC swabbing on stainless-steel plates, which were spiked with the diluted drug product. Different sampling and extraction techniques for TOC swabbing exist [12] and this study follows a procedure used in the pharmaceutical industry that is similar to those in other studies [16] for biopharmaceutical manufacturing and recommended by the PDA [1]. The drug product was diluted with purified water beforehand to ensure the desired TOC concentration. The swabbing procedure consists of using three swabs and the swabbing pattern can be seen in Fig. 2 in Supplemental material. The first two are wetted with purified water while the final swab is dry. After swabbing, 0.25 ml phosphoric acid is added to the TOC vials and the vials are filled with purified water to a final volume of 43 ml and then gently mixed by turning the vials upside down numerous times. The TOC samples are stored at 5°C C for up to three weeks until analysed. The stainless-steel plates are all cleaned in a laboratory dish washer with the same cleaning agents used in pharmaceutical production to ensure that insufficient cleaning does not affect the recovered TOC results [7].

Unless stated otherwise, each TOC sample was performed in triplicates, as recommended by the PDA [1]. Furthermore, each sample had two reference samples, which consisted of the same amount of the drug product that was spiked onto the stainless-steel plates, but directly deposited into two separate TOC vials. One vial was immediately filled with purified water and acid to be used as a check of the stock solution to ensure the proper concentration. The other vial was left to dry the same amount of time as the stainless-steel plates, in order to take any evaporation effects into account. This second reference sample is the one used when calculating the recoveries, which are calculated as [4].

\[
\text{recovery} = \frac{\text{TOC}_{\text{sample}}}{\text{TOC}_{\text{ref2}}} \times 100\% \tag{1}
\]

This is the approach used when the table value recovery factors are established, which are then used to calculate the final cleaning verification results. The area estimations were performed with blue dye on plastic sheets and in a tank that is no longer used for production. A high-resolution scan of the sheets or tank positions was then analysed with image analysis as described above. All area estimations and drug product recoveries were performed by trained personnel in the TOC swab procedure.

3. Results and Discussion

The TOC swab method itself is straightforward. Assuming a certain surface concentration, in µg/cm², the analyst swabs a specified area, e.g., 100 cm². This gives a total amount of residue on the swabs in µg, which is then extracted in a given amount of solvent, e.g., 43 ml, yielding a concentration in weight/volume, e.g., ppm (µg/ml) [1]. The resulting sample is analysed for TOC content. However, it is unlikely to swab every µg from the surface, which therefore necessitates recovery factors. A recovery factor is a correction factor used to relate the result in the vial to the amount that was on the surface. Therefore, the recovery factors must be established prior to the actual cleaning verification. Recovery factors are compound and surface type dependent [16], and generally there is one recovery factor per compound and surface type. All in all, the TOC result can thus be written with the formula

\[
\text{TOC [ppm]} = \frac{\text{area \left[ cm^2 \right]} \times \text{concentration \left[ cm^2/cm^2 \right]} \times \text{recovery}}{\text{volume}} \tag{2}
\]

Since area is to be determined and we assume the solvent volume to have a negligible error, two significant contributors to sampling error remain: recovery and area.

3.1. Recovery

As mentioned, this study focuses on 4 variables that influence the recovery: person dependency, position, surface finish, and concentration.

The person dependency of recovery results is well-documented, as described in other studies [3,4]. Data from eight people swabbing three plates each were used in this study.

It was then investigated how the positioning of the plates swabbed influenced the recovery. This is to imitate how it would be in a real-life swabbing situation where different positions in a tank can be reached. The imitated positions can be seen in Fig. 3 in the Supplemental material. Two persons swabbed in the same three positions 5 times.

Furthermore, it was investigated how both the surface finish, and concentration of the given analyte influenced the recovery. The finish has values ranging from #4–#8, with lower values indicating rougher surfaces.
The experimental data for the recovery as a function of finish and concentration were fitted with individual functions. The finish data were fitted with a linear function using the polyfit function in Matlab [15]. The concentration was fitted with a logistic function using the fit logistic [17] function of Matlab.

The results of both fits can be seen in the surface plot in Fig. 1. The figure shows that the recovery percentage drastically decreases towards lower values of concentration and that the recovery also decreases with decreasing surface finish. Thus, if measuring a very low concentration combined with a rough surface, one might risk having an actual recovery of 50% instead of 90% at high concentration and high finish. This is an issue as the actual recovery values used for correction, are usually determined at high concentrations and high finish. The different variables are assumed to be independent of each other.

Combining both the variability from different persons, from swabbing different positions, surface finish and concentration, it is possible to do a Monte Carlo simulation [18] of just how much the recovery can vary for a given product. This simulation is based on the individual relationships between recovery and the variables, which have been found by the experimental data. Since position and persons are discrete variables with no logical order, no fits were made for these variables.

The simulation begins with simulating the variance due to different analysts. This is done by having a recovery randomly chosen from a normal distribution with the mean and standard deviation being the same as the experimental distribution of recoveries from all persons.

Furthermore, a random scalar between 4 and 8 is chosen from a uniform distribution for the finish and used as the dependent variable in the linear function between recovery and finish. Then a random scalar between 0 and 4.5 is chosen from a uniform distribution. This is used as the independent variables in the logistic function describing the relationship between concentration and recovery. Uniform distributions are chosen as there does not exist data on the probability of surface finish in a random tank as well as the concentration levels in the tanks that are not biased by the sampling.

A random integer between 1 and 3 is chosen, indicating which position is used. If the position is not 1, the recovery will be corrected with regards to this factor.

The procedure is then repeated 1000 times and the histogram of the distributed recoveries can be seen in Fig. 2. Most recoveries are below 91%, which is the table value recovery. Acceptable RSD on recoveries can be found in the literature down to 15% [1,18]. 37.5% of the recoveries are below 76% recovery, i.e., 15% or lower than the table value, meaning 37.5% of the recoveries are outside the acceptable range defined by the RSD.

Combining both the variability from different persons, from swabbing different positions, surface finish and concentration, it is possible to do a Monte Carlo simulation [18] of just how much the recovery can vary for a given product. This simulation is based on the individual relationships between recovery and the variables, which have been found by the experimental data. Since position and persons are discrete variables with no logical order, no fits were made for these variables.

The simulation begins with simulating the variance due to different analysts. This is done by having a recovery randomly chosen from a normal distribution with the mean and standard deviation being the same as the experimental distribution of recoveries from all persons.

Furthermore, a random scalar between 4 and 8 is chosen from a uniform distribution for the finish and used as the dependent variable in the linear function between recovery and finish. Then a random scalar between 0 and 4.5 is chosen from a uniform distribution. This is used as the independent variables in the logistic function describing the relationship between concentration and recovery. Uniform distributions are chosen as there does not exist data on the probability of surface finish in a random tank as well as the concentration levels in the tanks that are not biased by the sampling.

A random integer between 1 and 3 is chosen, indicating which position is used. If the position is not 1, the recovery will be corrected with regards to this factor.

The procedure is then repeated 1000 times and the histogram of the distributed recoveries can be seen in Fig. 2. Most recoveries are below 91%, which is the table value recovery. Acceptable RSD on recoveries can be found in the literature down to 15% [1,18]. 37.5% of the recoveries are below 76% recovery, i.e., 15% or lower than the table value, meaning 37.5% of the recoveries are outside the acceptable range defined by the RSD.

Combining both the variability from different persons, from swabbing different positions, surface finish and concentration, it is possible to do a Monte Carlo simulation [18] of just how much the recovery can vary for a given product. This simulation is based on the individual relationships between recovery and the variables, which have been found by the experimental data. Since position and persons are discrete variables with no logical order, no fits were made for these variables.

The simulation begins with simulating the variance due to different analysts. This is done by having a recovery randomly chosen from a normal distribution with the mean and standard deviation being the same as the experimental distribution of recoveries from all persons.

Furthermore, a random scalar between 4 and 8 is chosen from a uniform distribution for the finish and used as the dependent variable in the linear function between recovery and finish. Then a random scalar between 0 and 4.5 is chosen from a uniform distribution. This is used as the independent variables in the logistic function describing the relationship between concentration and recovery. Uniform distributions are chosen as there does not exist data on the probability of surface finish in a random tank as well as the concentration levels in the tanks that are not biased by the sampling.

A random integer between 1 and 3 is chosen, indicating which position is used. If the position is not 1, the recovery will be corrected with regards to this factor.

The procedure is then repeated 1000 times and the histogram of the distributed recoveries can be seen in Fig. 2. Most recoveries are below 91%, which is the table value recovery. Acceptable RSD on recoveries can be found in the literature down to 15% [1,18]. 37.5% of the recoveries are below 76% recovery, i.e., 15% or lower than the table value, meaning 37.5% of the recoveries are outside the acceptable range defined by the RSD.

### 3.2. Area Estimation

In Eq. (2), the other important parameter is the area. As mentioned, it is assumed to be a constant value, however it is dependent upon the person performing the swabbing, and therefore a potential source of error. The results from the estimations of area, performed as described in the methods section, can be seen in Fig. 4 in the Supplemental material. The mean intrapersonal standard deviation is 13%, while the interpersonal standard deviation based on the mean recoveries for each person is 30%. Thus, while there is intrapersonal variance, the largest variance is related to interpersonal variances. Accounting for both these values, a potential difference of almost 100 cm$^2$ exists, i.e., the actual target area to be swabbed.

Furthermore, the area estimations in a tank were investigated, which can be seen in Fig. 5 in the Supplemental material. From the figure it is seen that these estimations are larger and up to 6 times greater than what they should be.

### 3.3. Simulation of TOC results

Combining the results from both the recovery and area leads to the simulation of how TOC results will be distributed when this sampling error is considered. A new Monte Carlo simulation is performed for each concentration, meaning that for a given concentration, a random person recovery, area, finish and position is chosen. There is no clear correlation between recovery and area ($r^2 = 0.33$), thus it was not decided to pair the recoveries and area estimations for the same person. We want to be able to show how the TOC results vary if we do not assume the recovery and area to be constant, but instead include the variation of the recovery and area. Thus, in Eq. (2), the numerator is multiplied by the simulated recovery, representing the actual recovery. The recovery in the denominator is then the table value recovery. The area is chosen randomly from a normal distribution based on the experimental data. For each concentration the simulation is performed 1000 times. The results for the RAL using lab and tank area estimations can be seen in Fig. 3. The orange histogram in the figure is the expected distribution if the actual recovery is the same as the reported recovery and with a swabbed area of 100 cm$^2$. The width of the distribution is due to the analytical uncertainty from the TOC analyser of 2% as reported by the manufacturer [20]. The left panel in Fig. 3 shows the results using the areas estimated in the laboratory and the panel on the right uses the areas estimated in the tank. The distribution using the tank areas was so
large that some TOC results became negative. As this has no physical meaning, these results were set to the critical limit, which is determined by the maximum allowed level of TOC in purified water [21]. This accounts for the first bin in the histogram with higher counts compared to the second bin. The difference in x-axes should also be noted and can be seen from the size of the orange histograms, which are the same in the two figures.

3.4. Limit of Detection

The results call for a revision of the limit of detection of the TOC swab method, as the determination of LOD is directly related to the error of the method. The correct figure of merit to use is the root-mean-square-error (RMSE) as it takes both accuracy and precision into account, compared to a standard deviation, which only considers precision. The RMSE is given by the formula

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} (\text{TOC}_i - \text{TOC}_{\text{expected}})^2}{N}}
\]

where \( N \) is the number of samples. However, given that Eq. (2) includes a division with the recovery, which is expressed as a percentage, the RMSE increases with concentration, i.e., heteroscedasticity. This can be seen in Fig. 6 in the Supplemental material, where a linear regression has been fitted to the RMSE as a function of concentration. The International Union of Pure and Applied Chemistry (IUPAC) has a definition of the LOD, when there is heteroscedasticity in the system, but still assuming normality [22]. These definitions are given by:

\[
\text{L}_{\text{D}} = \text{L}_C + \lambda_{0.95} \cdot (\sigma_b + a \cdot \text{L}_{\text{D}})
\]

where \( \lambda_{0.95} \) is the 95% quantile in the standard normal distribution. The value for \( \lambda_{0.95} \) can be found in any table with z-values for the normal distribution to be \( \lambda_{0.95} = 1.645 \) [22]. \( \text{L}_C \) is the critical level and \( \sigma_b \) is the uncertainty for a blank sample. For the TOC swab, the critical level is the maximum allowed level of TOC in the purified water [21] used for diluting the samples. Therefore, the uncertainty for a blank sample is also determined at this level. Using Eq. (4) with the slope from the linear regression and the uncertainty at the critical level, this yields an LOD of 60% of the RAL when using the laboratory estimated areas. Due to the large slope of the fit to the data using the tank areas, the LOD becomes negative. Naturally this does not have any physical meaning, and IUPAC also notes that this can occur in situations with large errors [22]. Considering the data for the LOD for the laboratory area estimations, this LOD is 2.4 times higher than the previous limit of detection, which was assumed to be equal to \( \text{L}_C \), as the analytical uncertainty on the instrument is very low (0.03 ppb) [20].

3.5. Identification of most critical variables

The interesting question is how to mitigate these findings? From the results and Eq. (2) the most influential factor for the TOC swab result is the area, as this is highly overestimated and up to 6 times higher than assumed. No single variable for the recovery has that significant of an influence. A simple solution to the overestimation of area could be to have area templates of stainless steel to be used in the tanks. However, depending on the size of the tank and position in the tank, the curvature of the tank surface might not be the same and to avoid cross-contamination one might need a wide variety of templates. Furthermore, cleaning verification positions are typically chosen to be the difficult-to-clean positions, such as under an impeller [1]. This presents a large challenge to make a template as the operator is swabbing blind. It
is not known if the overestimation of areas is proportionally large if using smaller areas, e.g., 25 cm². Forsyth [19] speculates that “it is also less likely to require the need for a template to take a 25 cm² sample”, which would reduce the issue with cross-contamination. However, there is no published data on this, thus further work should investigate whether the area is also overestimated at smaller target areas. Nonetheless, overestimation of the area will lead to high TOC swab results, i.e., ‘false positive’, where it will be mistaken that the equipment was not clean. This can have big economic impact for the pharmaceutical company due to scrapping of a batch that would not pose a risk for the patient. Instead, the assumption that recoveries are constant is riskier when they can be half as low as assumed. This will lead one to think that the surface concentration is below the RAL, when it is not, because only half of the contaminants were recovered compared to the recovery factor. Therefore it was simulated how the recovery factor would change when leaving out the variance in each variable one at a time. The results can be seen in Fig. 4, where calculated $\chi^2$ values between the distribution of recovery values that takes all factors into account and the distribution with the given variable variance left out. Only for the distribution without concentration variance, the $\chi^2$ value is above the critical value for a 5% significance level and thus the assumption that the blue and red distributions are the same, can be rejected. In other words, this variable is the most influential on the recovery percentage for this given drug product. Some debate exists as to whether the cleaning verification studies should be performed at a spiked level equal to the RAL [12] or the limit of detection/quantification (LOD/LOQ) [13]. In this study, these two are very similar and are thus both covered. Furthermore, there are opposing views as to whether it is useful to look at different spiked levels. The point of view of the PDA is “While it is possible to perform recoveries at different spiked levels, in general, there is little value to such additional spiked levels because of the variability of the sampling procedure” [12], while others argue that “the recovery factor derived from such a study may not accurately reflect the recovery for levels lower or higher than the level considered for the study, thereby possibly yielding false-negative or false-positive results” [22]. This study agrees with the latter view, as the data shows that influence of the concentration is statistically significant on the variance of the recovery. One way to mitigate the concentration level is to find the recovery factor at different concentrations and in order to safeguard always use the lowest value, even though this will also lead to overestimation in some cases. However, this worst-case approach has also been challenged [22], as any new lower outlier will change the recovery factor. Furthermore, PDA suggests that to qualify a sampling method, recoveries above 70% should be used, while in order to correct an analytical result, the recovery should be above 50% [12]. However, there is no scientific rationale for these limits [22] and this is an important discussion that needs to be taken within the pharmaceutical industry with input from the scientific community and regulatory bodies. The results of this study are only for this specific large molecule drug product and the extension to the TOC swab results for other API’s and products should always be investigated separately. Instead, this study is meant as an eye-opener for the sampling bias in TOC swabbing. Furthermore, and most of all, it is an invitation to the rest of the pharmaceutical industry to start this discussion and investigation of the sampling error in TOC swab. This paper has shown that with relatively simple experiments and simulation, it is possible to give a quantitative estimate of the sampling error and how to use the simulation to figure out which factors are most important to mitigate.

4. Conclusion

This paper quantified the sampling error in TOC swab by using Monte Carlo simulation for a given drug product, when considering different variables that will occur in an actual cleaning verification setting. These variables were variance due to surface finish, residue concentration, swabbing position and different persons swabbing, which influences both the recovery and the area swabbed. Including the variance from all these variables lead to a highly increased LOD 2.4 times higher than the current LOD when using areas estimated in the laboratory compared to the LOD equivalent to the critical limit when only considering the analytical TOC method. If using the areas estimated in a tank, the large heteroscedasticity meant that an LOD with physical meaning could not be established.

Fig. 4. Distribution of recoveries when removing different factors to test which factor alters the distributions most. Upper left has no variation from concentration, upper right has no variation from persons, lower left has no contribution from finish and lower right has no contribution from position. Only upper left plot has a $\chi^2$-value higher than the critical $\chi^2$-value at the given significance level of 0.05 and degrees of freedom.
Furthermore, it was investigated which of these factors is most influential on the large sampling error. It was found that the area is grossly overestimated, leading to artificially high TOC results. The second most important factor was the concentration, as the recovery factor was found to drastically decrease with decreasing concentration for the given product. Based on these results, this paper is an invitation to discuss and investigate the sampling error of TOC swab.

CRediT authorship contribution statement

Isabella Jul-Jørgensen: Writing – original draft, Methodology, Formal analysis, Investigation, Visualization, Conceptualization. Christian A. Hundahl: Writing – review & editing, Methodology, Conceptualization, Supervision. Erik Skibsted: Writing – review & editing, Conceptualization. Krist V. Gernaey: Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Funded by Novo Nordisk A/S and the Innovation Fund Denmark with Grant no. 0153-00090B. The authors would also like to acknowledge the assistance of Christina Leje Simonson and Pernille Beicker Petersen for laboratory assistance in the experimental part. Furthermore, we would like to acknowledge Joachim Brink Harck for writing the image analysis script used to analyse the area estimations.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jpba.2022.114763.

References