Chemosphere 300 (2022) 134476

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Unravelling reasons for variability in the OECD 306 marine biodegradation test

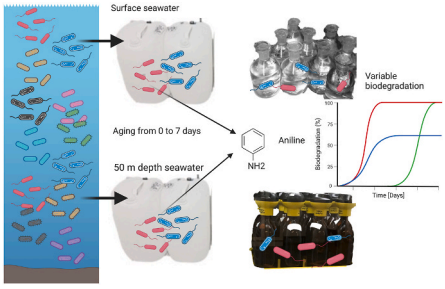
Aina Charlotte Wennberg [a,*], Sondre Meland [a], Merete Grung [a], Adam Lillicrap [a]

[a] Norwegian Institute for Water Research (NIVA), Økernveien 94, NO-0579, Oslo, Norway

HIGHLIGHTS

- Marine biodegradation test results according to OECD TG306 are variable.
- Biodegradation of aniline varied significantly between sampling seasons and depths.
- Microbial community changed with time during aging and according to test set up.
- A more robust and consistent microbial inoculum source is needed to reduce variability.

GRAPHICAL ABSTRACT



ARTICLE INFO

Handling Editor: Willie Peijnenburg

Keywords:
Biodegradation
Marine environment
Chemical fate assessment
Seawater inoculum
Microbial community
OECD 306

ABSTRACT

The recommended test for assessing if a chemical can be biodegraded in the marine environment is performed according to the Organisation for Economic Cooperation and Development Marine biodegradation test guideline (OECD 306). However, this test is known to generate highly variable test results when comparing interlaboratory test results for the same compound. One reason can be the relatively low bacterial content compared to the inoculum used for OECD readily biodegradation tests (OECD 301). Some of the variability in data obtained from OECD 306 tests can also be due to the flexibility on how to store the seawater inoculum before starting a test. Another variable in the seawater inoculum is the source of seawater used by different laboratories, i.e., geographical location and anthropogenic activities at the source. In this study, the effect of aging seawater and the source of seawater (sample time and depth) were investigated to determine differences in the biodegradation of the reference compound aniline. Aging the seawater before starting the test is recommended in OECD 306 to reduce the background levels of organic carbon in the water. However, it also functions to acclimatize the bacterial community from the environmental source temperature to the test temperature (normally 20 °C). Herein, the microbial community was monitored using flowcytometer during the aging process. As expected, the microbial community changed over time. In one experiment, aging significantly improved the biodegradation of aniline, while in two experiments, there was no significant difference in biodegradation. Interestingly however, there was significant variability in the biodegradation of aniline between sampling seasons and depths, even when all experiments were performed in the same lab, by the same operator and seawater obtained from the

Abbreviations: BST, Biodegradation Screening Test; DO, Dissolved Oxygen; DOC, Dissolved Organic Carbon; FCM, Flow cytometer; PCA, Principal Component Analysis; pRDA, partial Redundancy Analysis; RDA, Redundancy Analysis; ThOD, Theoretical Oxygen Demand.

* Corresponding author.
E-mail address: aina.charlotte.wennberg@niva.no (A.C. Wennberg).

https://doi.org/10.1016/j.chemosphere.2022.134476
Received 4 January 2022; Received in revised form 26 March 2022; Accepted 28 March 2022
Available online 31 March 2022

same source. This highlights the need for a more robust and consistent microbial inoculum source to reduce variability in seawater biodegradation tests.

1. Introduction

A chemical's potential to degrade is a defining criterion when considering the risk of exposure in the aquatic environment. Several persistent chemical groups have proved to be problematic many years after their application (Cousins et al., 2019). Thus, persistence of a chemical is an important criterion for regulating use and release of chemicals. According to EU Regulation 1907/2006 of the European Parliament and of the Council of December 18, 2006 concerning the Registration, Evaluation, Authorisation and restriction of Chemicals (REACH), a tiered testing strategy can be used for a stepwise decision on whether a substance is not persistent, potentially persistent or persistent. The first step is the biodegradation screening tests (BST) for readily biodegradability. The recommended test guidelines for BSTs are developed and approved by the Organisation for Economic Cooperation and Development (OECD) and include OECD 301 test series (A to F) and OECD 310 which use bacterial inoculum from wastewater treatment plants (sludge or effluent) or surface freshwater sources (OECD, 1992a). In addition, there is a test for biodegradability in seawater (OECD 306). Besides REACH, OECD 306 is also recommended as one of the tests for assessment of offshore chemicals according to the convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR Convention). OSPAR was adopted by 15 European countries and contains decision points, recommendations, and agreements regarding regulations on the use of chemicals. To reduce the overall impact of offshore chemicals on the marine environment, OSPAR has implemented a Harmonized Offshore Chemical Notification Format (HOCNF) that chemical suppliers must submit for all chemicals used in connection with offshore exploration and production activities in the OSPAR maritime areas (OSPAR, 2012, 2015). The guidance document for HOCNF states that toxicity and biodegradation data should be obtained with standardized marine tests, (OSPAR, 2012).

Even if OECD 306 is a much used and recommended test guideline, there are some concerns that the OECD 306 test has high variability in test results, as discussed in a multistakeholder workshop (ECETOC, 2017; Ott et al., 2019). There are several known factors that can cause variability in biodegradation testing. The most important microbial factors that affect the results of biodegradation lab tests are the source, diversity and abundance of microorganisms in the test inoculum (Kowalczyk et al., 2015). Diversity of marine bacteria changes both with season and latitude and human environmental impacts (Ladau et al., 2013). The marine biodegradation test (OECD 306) often has a lower concentration of bacteria than the ready BST (e.g. OECD 301), making it a more conservative test for assessing persistence (ECETOC, 2017). Both increasing the concentration of bacteria or using larger test volumes of seawater inoculum has been demonstrated to increase the potential for biodegradation (Martin et al., 2017). Laboratories performing marine biodegradation tests depend on availability of a location for sampling fresh natural seawater. The laboratories that participated in an international ring test in 2017, used seawater from depths spanning from 0 to 60 m at locations from 40 to 5000 m from the shore across three continents (Ott et al., 2020a). In the ring test, the original OECD 306 test set-up with closed bottles had a coefficient of variance (CV) of 49% when including the negative control and assessing three non-persistent and one persistent test chemical. Changing the test set-up to respirometers and extending to 60 days incubation reduced CV to 42% and using seawater with 100 times nominal concentration of bacteria reduced it further to 35%.

Microorganisms use enzymes for degrading organic matter to be used as a source of energy or growth but not all bacteria have the necessary genome to produce all the enzymes necessary to break all types of chemical bonds. Hence, biodegradation is not only dependent on a sufficient number of bacteria to be present in the test but also a sufficient diversity (Davenport et al., 2022; Ott et al., 2020a).

Another factor that can affect the results of biodegradation lab tests is the time the bacteria require for adapting to a new environment or producing the necessary enzymes they need to biodegrade a substance. This is referred to as the lag-phase in biodegradation tests. Studies have shown that marine bacteria often exhibit longer lag-phases than freshwater bacteria (ECETOC, 2003a). Modifications in test set-up such as acclimatizing (or aging) the test inoculum or pre-exposure of bacteria to the test substance (adaption) can influence the lag-phase (Thouand et al., 1996). Aging the test water for up to one week is acceptable within the OECD 306 test guideline, while adaptation of inoculum to the test substance is not accepted for screening or simulation tests (ECHA, 2017).

Natural seawater contains a complex microbial community, with variability in species due to both geographical and seasonal differences, and can have different dynamics and resilience (Fuhrman et al., 2015). To be able to understand what happens to the microbial community in the seawater, when it is brought into the lab, a method is needed to analyze changes in the microbial community. The most accurate method would be to identify all bacteria by sequencing and extensive bioinformatics. However, sequencing requires both significant resources for DNA-extraction and molecular analysis, extensive computing power and highly trained bioinformatics. This causes limitations on the number of samples that are feasible to analyze within a reasonable budget (Lambrecht et al., 2018). An alternative method is the use of flowcytometry (FCM), which is a relatively quick and easy method to monitor changes in bacterial community in a water sample with little sample pre-treatment compared with sequencing studies (Buysschaert et al., 2018; Kinet et al., 2016; Prest et al., 2014). The bacterial community can then be detailed using a statistical method called fingerprinting, which can be used to quantify major changes in the bacterial community, even though it cannot identify the specific bacterial species present (Koch et al., 2013; Props et al., 2016). This has previously been demonstrated with bacteria obtained from landfill leachate (Kinet et al., 2016), waste water (Guenther et al., 2012), drinking water pipes (Van Nevel et al., 2017) and sub-arctic peat land (Quiroga et al., 2017). There are different strategies for setting up a bacterial fingerprint. The first step is the separation of the microbial community into distinct clusters or bins. This is dependent on the software used, the packages available in the statistical software R, and whether a strictly automatic and statistical method (Kinet et al., 2016; Koch et al., 2014; Quiroga et al., 2017; Van Nevel et al., 2016, 2017) or a manual method for identifying clusters of similar bacteria, is used (Koch et al., 2013, 2014). The second step is the comparison of samples by use of different statistical methods such as heat maps, dendrograms, principal component analysis (PCA) or dissimilarity matrixes.

The OECD 306 test guideline does not have any recommendation of the source of the seawater, only that sampling site and water quality should be reported related to pollution and nutrient status. Recommended pre-treatment is aging for approximately a week if the dissolved organic carbon (DOC) content is high, and to be either filtered or sedimented to remove particles (OECD, 1992b). In NIVA's laboratory, seawater is obtained from the Oslofjord through a permanent pumping installation taking in both surface water and deep water from approximately 50 m depth. For biodegradation testing we use seawater collected from 50 m depth based on the assumption that this water provides more consistent biodegradation results compared to surface water. The microbial community in the surface water can be affected by the variability of phytoplankton and zooplankton blooms from early

spring to late autumn, while the deep water has more stable conditions (none published data). Another concern for the test set-up is the difference in temperature between the fjord and the lab. Typical temperatures for the deep-water ranges between 6 °C and 9 °C throughout the year, while the biodegradation tests are performed at 20 °C. Thus, it was evaluated whether lowering the test temperature to 15 °C would give better degradation by the cold-adopted microbial community. The normal strategy by the laboratory is to store the seawater at the test temperature (20 °C) for a few days before the test set-up to give the microbial community a chance to acclimatize to the testing condition, assuming this might reduce the lag phase of the test and thus give more rapid biodegradation. Storing the seawater before the test, so called aging, is recommended in OECD 306 if the DOC concentration is high, however, the typical DOC content of our water is around 1 mg/L and aging is used as a means of acclimatization. It is known that the diversity of the microbial community is affected by laboratory conditions, and that storing the seawater before testing (aging) might change the microbial composition compared to the environment (ECETOC, 2003b). However, if this change in microbial composition is an effect of the laboratory conditions, the change will potentially still happen in the test bottles, and might lead to a longer lag-phase, if adaptation to laboratory conditions is not allowed to happen before the start of the test. In contrast, the presence of the test chemical, as a food source during acclimatization, could result in a different community structure with a higher probability of survival and growth of the competent degraders.

To better understand the variability in the OECD 306 test, the aim of this study was to assess if aging the seawater and using seawater collected from different depth would provide more reliable biodegradation results for the OECD 306 test set-up. In addition, reducing the test temperature from 20 °C to 15 °C was evaluated to see if a lower temperature would be better for the cold seawater microbial community. Evaluation of successful biodegradation was based on the level of aniline degraded after 7 days, as this reference compound has been seen to have a variable lag-phase of zero to seven days and time to achieve 50% degradation, excluding the lag phase, of two to twelve days (Nyholm and Kristensen, 1992). Experience from our laboratory is that aniline is typically degraded between 57% and 76% on day 7 (non-published data). A second aim was to improve knowledge on how the microbial community respond to laboratory conditions before and after starting biodegradation tests.

2. Materials and methods

2.1. Experimental design

A first experiment was preformed to answer the questions: 1. Is aniline degrading faster in seawater from 50 m depth than surface water? 2. Is biodegradation of aniline faster after aging/acclimating water at the test temperature? 3. Will aniline degrade as fast at 15 °C as at 20 °C when using seawater samples at colder temperatures? 4. For test set-up convenience, can mineral nutrients be added to the test water before aging? 5. What causes the most change in microbial diversity of the above-mentioned tests?

To address questions 1 to 4, seawater was added with aniline and poured into bottles and analysed for dissolved oxygen before sealing. Dissolved oxygen was also measured after 7 days and compared to blank controls consisting of seawater without aniline. To answer question 5, samples were taken from storage cans, test bottles and blank control bottles, before and after incubation, and analysed using FCM.

Due to the number of questions to be answered, limited resources and time, only one biological replicate was included in experiment one.

The second experiment was performed using three biological replicates repeating experiment one to give more data for answering question 1, 2, and 5.

In experiments one and two, the success of aniline degradation was only assessed based on one datapoint (day 7), thus it was not possible to assess lag-phase and total biodegradation. To improve this, a third experiment was performed to give more information on the degradation curves for aniline using respirometer and increasing to a full incubation duration of 28 days. The respirometers could not be opened for sampling for analysis of microbial community without affecting the results, thus additional bottles for blank controls were included so that experiment three could also provide more data on assessing changes in microbial community.

Repeating the experiment three times meant that water was collected on three different times of the year from November to April and enabled an investigation into whether a seasonal change might also affect the results. Thus, the seasonal effect also needed consideration in interpretation of the results.

2.2. Seawater collection

Seawater was collected from the Oslofjord at NIVA's marine field station in Solbergstrand, Drøbak, Norway (59.61574° N, 10.65280° E) at three different times: November 2016, February 2018 and April 2018. The water was pumped from either surface (1 m depth) or deep water (50 m depth) through a permanent installation and filled into cleaned (mild chlorination followed by rinsing with tap-water) 25 L plastic jerrycans. The water was transported to the NIVA laboratory in Oslo within 2 h and aged at test temperature covered with dark plastic to protect from light.

2.3. Experimental set-up and sampling

Three experiments were performed in November 2016, February 2018 and April 2018 with slight differences in test set-up, sampling times and analysis. An overview of the test set-up including number of replicates, incubation times and type of analysis are given in Table 1.

3.3.1. Test chemical

Aniline (CAS: 62-53-3, ≥99,5% purity, Sigma Aldrich) is used as a reference substance for OECD 306 biodegradation tests. Stock solutions were prepared at 2 g/L in MilliQ water and added to seawater test medium at a final concentration of 2 mg/L (ThOD 4.9 mg O_2/L) for the closed bottle test set-up and a 4 g/L stock solution, and final concentration of 33 mg/L (ThOD 80 mg O_2/L), for the respirometer test set-up.

3.3.2. Test set-up

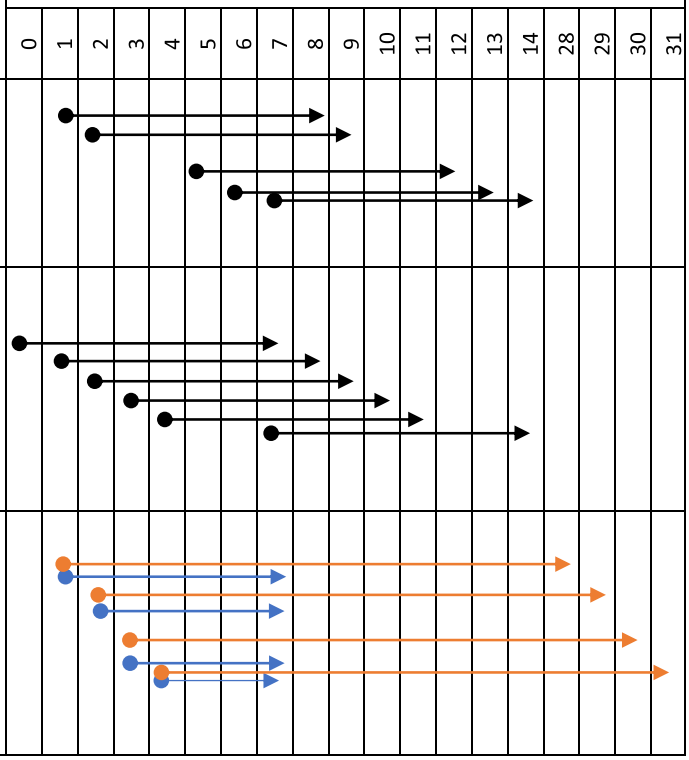
290 ml glass bottles with ground glass stoppers were used for the closed bottle test set-up according to OECD 306. Bottles were filled with seawater medium supplemented with mineral supplements according to OECD 306 with or without aniline until overflowing and then stoppered ensuring no head space. The bottles were incubated for 7 days in a climate-controlled room covered with black plastic to protect from light. Biodegradation was measured as reduction in dissolved oxygen corrected for blank respiration. For continuous monitoring of biodegradation curves in experiment 3, respirometers (WTW OxiTop®-C) were used according to OECD 301 F (OECD, 1992a), however, with mineral nutrient supplements as specified in OECD 306. Bottles were filled with 250 ml test medium, leaving a head space of 250 ml. Respirometers were incubated for 28 days in the dark with continuous stirring and NaOH added as an adsorbent for CO_2. No sampling was performed from the respirometers, but a set with closed bottles for blank controls were included for analysis of microbial community.

3.3.3. Optimizing test set-up

In the first experimental round performed in November 2016, the experiment included two temperatures (20 °C and 15 °C) and an evaluation of the effects of aging the seawater with and without mineral nutrients. The results of this experiment are shown in the supplementary material. The remaining experiments were performed at 20 °C and with mineral nutrients added before storage, and only data from these test

Table 1
Experimental set-up, temperature and salinity of seawater at collection time and incubation period for biodegradation set up. Sampling for microbial community analysis by FCM was performed with 1–3 biological replicates from storage cans and blank controls on days with incubation start and stop (black arrow) for experiment 1 and 2, and on incubation start and day 7 (blue arrow) for experiment 3. Biodegradation was evaluated with either three closed bottles (black arrows) or two respirometer bottles (orange arrows). Sampling and analysis were not done during weekends, thus the gap in sampling times.

Date/ set-up	Surface water	Deep water	Replicates	Days after seawater collection
Experiment 1 Nov. 9, 2016 / Closed bottle	7.0 ˚C, 32 PSU	7.5 ˚C, 34 PSU	1 FCM 3 biodeg	
Experiment 2 Feb. 19, 2018/ Closed bottle	1.6 ˚C, 25 PSU	8.4 ˚C, 35 PSU	3 FCM 3 biodeg	
Experiment 3 April 17, 2018/ respirometers	5.8 ˚C, 20 PSU	7.2 ˚C, 35 PSU	3 FCM 2 biodeg	



set-ups are included in the following analyses.

3.3.4. *Effect of laboratory conditions on microbial community*

Samples were taken from the storage cans during aging on every weekday for analysis of bacterial counts and community structure using FCM. In addition, samples were analysed in blank controls filled at different days of aging after incubation for 7 days and in test bottles with aniline in experiment 1 and 2.

2.4. *Analysis*

Dissolved oxygen (DO) was analysed with an electrochemical oxygen probe (WTW Inolab Oxi 7310) in all closed bottles after 7 days incubation for experiments 1 and 2 (November and February). Consumption of oxygen was measured continuously for 28 days using OxiTop®-C measuring heads with sodium hydroxide as adsorbent of produced CO_2 in experiment 3 (April).

Samples for the microbial count were analysed the same day of sampling, without any preservation, using FCM. Samples were diluted 1:10 in deionized water (MilliQ) and stained with SYBRGreen I (x10000 concentration in DMSO (Invitrogen), working stock 1:100 diluted in Tris-Buffer, 1:100 diluted in sample) at room temperature for a minimum of 20 min. SYBRGreen I binds to DNA in the bacterial cell, and is routinely used for bacterial total counts (Hammes et al., 2008). The volume used for analysis of each diluted and stained sample was 50 μl and were run on an Accuri C6 (BD) flow cytometer fitted with a blue laser (488 nm), and the detectors for green (533/30 nm) and red (>670 nm) fluorescence were used for analysing the data in the BD Accuri™ C6 Software. Bacteria were separated from background noise using electronic gating (Prest et al., 2013).

2.5. *Data handling and statistics*

Oxygen measurements and OxiTop® data were imported to Excel and the percentage biodegradation was calculated based on the theoretical oxygen demands according to OECD 301 test guideline. Graphs were created, and statistics performed using JMP®16 (SAS Institute Inc.). Comparison of biodegradation results as response to the parameters sampling time (November, February and April), depth (Surface and Deep) and aging time was evaluated using a general linear model (GLM) with least squares fit using linear regression and significance level $p \leq 0.05$. Each experiment was analysed separately. Interaction between depth and aging was investigated with no effect. Therefore, the model was constructed with only depth and aging alone, not including the combination of these. To test if there were differences between results from the different experiments (November, February and April), an Anova test followed by a Tukey post-hoc test was performed.

FCM data was collected using the BD Accuri C6 Software and imported to R using Bioconductor and the FlowCore package (Ellis et al., 2020). A fingerprint was made by creating 32 mathematical bins using the FlowFP package (Rogers and Holyst, 2009) and compared using flowCyBar (Koch et al., 2013). A description of the workflow for the data handling in R is included in supplementary text S.3 Workflow in R for fingerprint analysis of bacterial community.

Univariate and multivariate statistical analysis of bacterial fingerprints were calculated using Canoco 5.12 software (ter Braak and Smilauer, 2018). To assess the variation in the multivariate response variables "bacterial fingerprints", the ordination techniques principal component analysis (PCA), redundancy analysis (RDA) and partial RDA (pRDA) were applied. PCA is an unconstrained method that assesses the maximum variation in the observed data, whereas RDA is a constrained method that assesses how much of the observed variation can be ascribed by one or several explanatory variables. Like RDA, pRDA assesses any relationship between the observed data and the explanatory variable(s). However, in the pRDA the effects of covariates are removed before RDA is performed on the remaining variation. Hence, by using pRDA it is possible to reveal the unique effect of each explanatory variable or group of variables. The relationship between the observed variation and the explanatory variable(s) were tested using Monte Carlo permutation tests (1999 permutations) and the level of statistical significance was set to $p < 0.05$.

3. Results and discussion

3.1. Optimizing test set-up

In the first experiment performed in November 2016, both a variation of incubation temperature and time of addition of mineral nutrients indicated that 7 days was not sufficient to ensure degradation of aniline at 15 °C even if the microbial inoculum was collected at a sea temperature of 7 °C, see supplementary results S1. This is not unexpected, as both chemical reactions and biological processes are slower at colder temperatures. It is recommended to consider the source of the inoculum when choosing the test temperature, however, Sjøholm et al. (2022) found that heating a winter sample to test temperature gave less influence of the biodegradation rate than cooling a summer sample to winter temperatures.

Adding the mineral nutrients in the storage tanks upon arrival to the lab instead of adding nutrients after aging was convenient for the experimental test set-up since it reduced workload. The time of addition of nutrients did not affect biodegradation results (see supplementary results S2), thus experiments 2 and 3 were carried out with mineral nutrients added before aging.

3.2. Comparing biodegradation to aging time and sampling depth

The main questions for this study were: Is aniline degrading faster in deep seawater than surface water? And is biodegradation of aniline faster after aging/acclimating water at the test temperature?

The level of biodegradation of aniline after seven days of incubation in all three experiments are presented in Fig. 1.

There was no significant difference between biodegradation of aniline in deep seawater compared to surface water in the 1st test (November experiment ($p = 0.6$, mean 49% for deep and 45% for surface)), while biodegradation was higher in surface water compared to deep water in the second test (February ($p = 0.00001$, mean 46% for deep and 64% for surface)) and was higher in deep water than surface water in the 3rd test (April ($p = 0.005$, mean 68% for deep and 24% for surface)). Comparing the deep-water results from the three experiments, the 3rd experiment had significantly higher biodegradation of aniline (71%) than in the 1st and 2nd experiments (45% and 48%, respectively; $p = 0.008$). However, when comparing the surface water results, the 2nd experiment (February) had significantly higher biodegradation (65%) than the 1st and 3rd experiments (November (43%) and April (26%), $p < 0.0001$).

Aging the water only had a significant effect on biodegradation of aniline in the 1st (November) experiment, where there was a positive effect of aging on the biodegradation of aniline on seawater from both depths ($p = 0.0005$, slope 6% per day, std. error 1.5%). Although not statistically significant, there also seemed to be a positive effect from aging the water on the deep water sampled in April (exp 3, $p = 0.22$ slope 7%, std. error 5.8%). In the rest of the samples, neither a positive nor negative effect of aging the water was seen on the biodegradation of aniline.

Aging seemed to improve the reproducibility of the biodegradation in parallel bottles only in the 1st test (November) in the surface water sample where the standard deviation was significantly reduced with increasing aging time ($p = 0.0033$, slope – 6.5% std. error 0.2%), while no significant trend was seen in the other experiments.

November and February (1st and 2nd experiments) samples were run in closed bottles with an aniline concentration of 2 mg/L and had little variation between triplicate bottles, except for bottles started after 1 and 2 days aging in November and without aging in February. The April

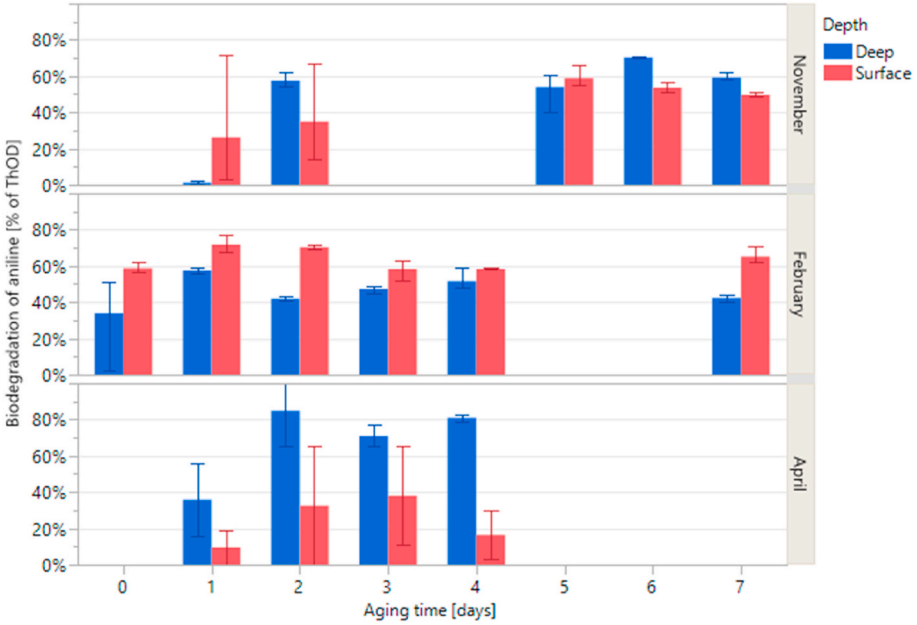


Fig. 1. Biodegradation of aniline after incubation for 7 days at 20 °C after aging the seawater for 0–7 days. Each error bar is constructed using the minimum and maximum of the data (three data points for November and February, two for April).

experiment was performed with OxiTop® bottles with an aniline concentration of 33 mg/L and had a larger variance between duplicate bottles. This could be due to either the higher test concentration, the technical conditions of the OxiTop® heads or the difference in the microbial community.

To determine whether degradation of aniline is faster with aging, the endpoints, lag phase and time to reach 60% biodegradation was included in addition to biodegradation on day 7 for the 3rd experiment (April) using OxiTop®-bottles. These benchmarks are compared in Fig. 2. It seemed that the samples aged for two days had faster degradation of aniline compared to one day: both lag time and time to reach 60% degradation was shorter, but on average it increased when aging for three and four days. However, with only four datapoints for each aging time and large variations in parallels, it is not enough to conclude on any trends.

This study is a combination of three experiments with slight differences in experimental design, the major difference being the use of closed bottles and the respirometer set-up, with a difference in initial test concentration and method of DO measurement. However, since the conclusions of this study are based on comparisons within each experiment and not between experiments, the differences in test set-up are considered of less importance. The answers to the research questions are, however, ambiguous.

Using seawater collected from 50 m depth compared to surface water did not give better degradation results in all experiments, and not less variability either. There seem to be a tendency towards faster biodegradation and more reproducible results with aging seawater, however, there is still a large variation in results depending on seawater collection time and depth.

These experiments were performed with an incubation time of only 7 days to have a timepoint within the typical exponential biodegradation curve for aniline. In a ring test (Nyholm and Kristensen, 1992), the typical degradation of aniline had a lag phase of between 0 and 7 days and time to achieve 50% degradation, not considering lag-phase, was between 2 and 12 days (Nyholm and Kristensen, 1992). Thus, the results presented here are not atypical. However, they demonstrate how much variation in experimental results can be expected even when removing the variables of different laboratories, different water source and different testing personnel. From all experiments performed herein, five surface water samples and four deep water samples failed to pass the lag phase within seven days.

Even if this study was only performed with one chemical, it is likely that a variation would also be seen with other chemicals related to how fast they degrade. A more rapidly degrading chemical might have a smaller variance, while a more difficult to degrade chemical might have a larger variance as it might be more challenging to capture a competent microbial community in the test. However, if the same conditions would give faster or slower degradation for all chemicals cannot be speculated based on the results of only the one test chemical aniline.

3.3. How laboratory conditions affect microbial community

Microbial community was analysed using FCM, and the difference over time in fingerprint of the community of bacteria in seawater was compared using PCA and RDA. Seawater was either stored only in storage cans ("can") from day 0 to day 14 or filled from the storage cans at different aging times (day 0–7) and incubated as blank controls ("bottles") for 7 days and analysed from day 7 to day 14, except for experiment 3, where they were all analysed on day 7. Time had the largest effect on the observed variation in the microbial community, explaining up to 84% of the variance seen in the different samples from cans and blank controls samples from day one to day 14 after collecting seawater. Whether the water was only stored in the 20 L jerry cans or incubated in the small 290 ml glass, explained the variation in the microbial community in the range of 0–38%, see results for "vessel" in Table 2. This showed that the microbial community in seawater will change over time from the original community collected in the natural environment as has been seen in other studies (ECETOC, 2003b). However, some of the variation can also be attributed to the transfer to the bottles used for the blank control and test bottles. Thus, any change in external factors might change the microbial community in the laboratory.

Liu et al. (2019) described how natural microbial communities can be studied using flow cytometer to analyze communities and their dynamic behavior. They utilized the stability properties resistance, resilience, displacement speed, and elasticity normally applied to ecosystems on microbial community data from FCM. One of their observations was that parallel samples of complex microbial samples did not change at the

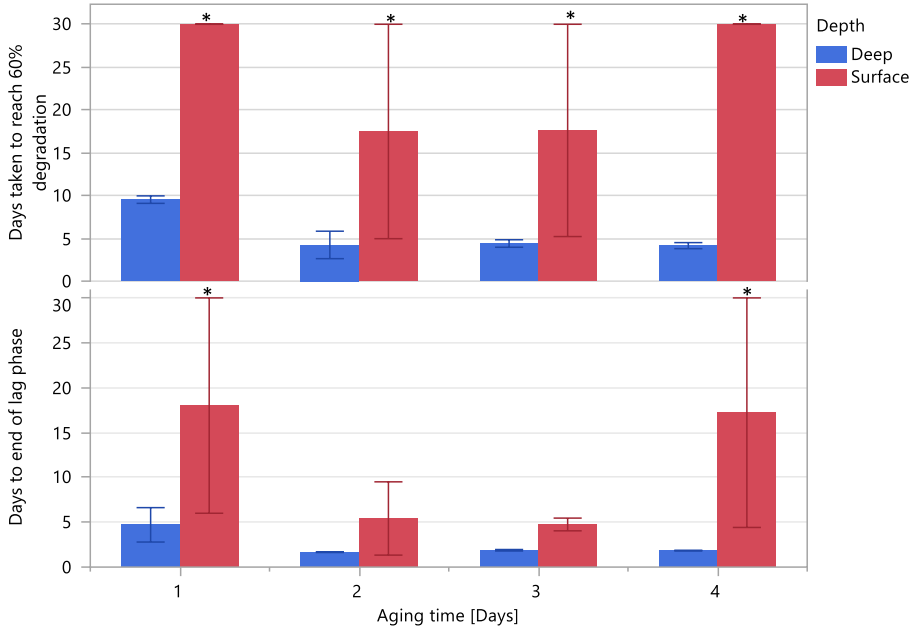


Fig. 2. Biodegradation of aniline in OxiTop® bottles: End of lag phase defined as 10% biodegradation. Each error bar is constructed using the minimum and maximum of the two data points. * not reached within 28 days incubation time.

Table 2
The output of the RDA and pRDA assessing the variation in the microbial community in different months and depth that can be ascribed to the explanatory variables, vessel type and aging time (included as factors (i.e. nominal variable) and not numerical variable). Statistical analysis of the effect of storage conditions on variation in microbial community measured by change in fingerprint of FCM analysis. Green cells are statistically significant results, while red cells are not significant.

Sampling time	Sampling depth	Number of samples	Constrained RDA [% adjusted explained variation]		Variation partitioning with pRDA [% variation explained] a + b + c = total effect, a = unique effect of vessel (time as co-factor), b = unique effect of time (vessel as co-factor), c = joint effect			
			Vessel (Can vs. bottle)	Time	Vessel + time (a+b + c)	Vessel only (a)	Time only (b)	Joint explanation vessel + time (c)
November	surface	13[a]	7.4 p = 0.06	0 p = 0.6	47 p = 0.08	53 p = 0.06	40 p = 0.1	−46
	deep	13[a]	3.3 p = 0.2	33 p = 0.3	64 p = 0.1	31 p = 0.08	60 p = 0.2	−28
February	surface	37[b]	8.8 p = 0.008	74 p = 0.001	82 p = 0.001	8.1 p = 0.001	74 p = 0.001	0.7
	deep	37[b]	0.99 p = 0.2	84 p = 0.001	86 p = 0.001	1.8 p = 0.002	85 p = 0.001	−0.8
April	surface	51[c]	22 p = 0.001	83 p = 0.001	84 p = 0.001	<0.1 p = 0.2	62 p = 0.001	22
	deep	54[c]	38 p = 0.001	48 p = 0.001	53 p = 0.001	5.3 p = 0.001	15 p = 0.001	33

[a] Sampling storage cans on day 2, 5–9, 12 and 13 and blank controls on day 9, 12, 13 and 14, no replicates.
[b] Sampling storage cans on day 1–4, 7–11 and 14 and blank controls on day 8–11 and 14, analytical duplicates or triplicates.
[c] Sampling triplicate storage cans on day 1–4 and 7 and triplicate blank controls on day 7, analytical duplicates. For surface water three deviating analysis were excluded because of analytical error.

same time and in the same way (Liu et al., 2019). This might explain some of the variability in the observations in the experiments herein, as bottle to bottle variation is also an important factor in evaluating these results. Only single replicate vessels were analysed in the November and February water samples, while triplicates were analysed on the April water samples. To better understand what is driving this change, a more detailed study of the seawater inoculum could be useful by implementing a more time intensive sampling regime at the critical time from sampling, transport and first acclimatization. This would enable the possibility to map the bacterial ecology in terms of resilience, resistance and elasticity (Liu et al., 2018), and could perhaps be used to compare different seawater sources.

3.4. Bacterial abundance during storage and testing

Both November and February samples showed an increase in bacterial count during the first days of storage, followed by a slight decrease. In both cases, the increase was more rapid and larger in surface water than deep water. However, the April samples had a slight decrease in bacterial counts. These numbers represent the abundances of pelagic bacteria (free bacteria in the water phase), not including bacterial flocs or biofilms. It is likely that bacteria would also form biofilm on the surface of the cans, thus, the total biomass cannot be estimated based on FCM counts of the water phase alone. At all sampling times, the surface water had slightly higher bacterial abundance than deep water sampled at the same time, see Fig. 3, and is within the normal level for this test of 10^5-10^7 cells/ml (Davenport et al., 2022). Bacterial abundance is one of the factors influencing the probability for a compound to

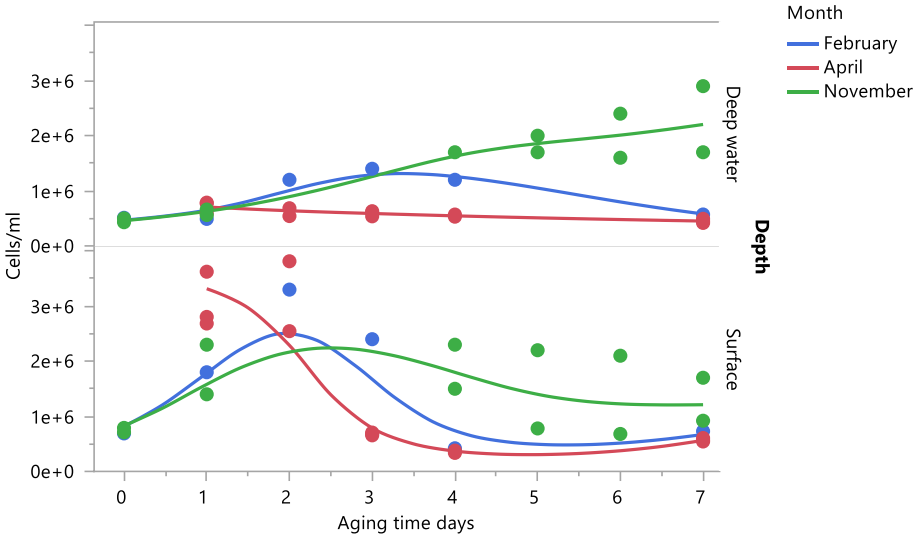


Fig. 3. Bacterial count between day 0 and 7 in storage cans filled with surface or deep water collected in November (two containers), February (one container) and April (three containers). Samples where not taken during weekend.

be biodegraded. Therefore, the rational for increasing the bacterial abundance, by concentrating water for the enhanced OECD 306 screening test in the international ring test trial, was that increasing the bacterial abundance from the raw water would also increase the biodiversity (Ott et al., 2020a, 2020b). Comparing bacterial count and aniline degradation in this study did not show any correlation (results not shown).

In November and February experiments, samples from test bottles with aniline were also analysed by FCM and revealed both increased bacterial abundance and larger particles that were either caused by bacterial flocks and/or protozoa which was confirmed by microscopy. The frequency of increased bacterial abundance and/or larger particles was higher for bottles where aniline was degraded than in bottles which failed the degradation criteria or in the blank controls. Biodegradation screening tests, where the test chemicals are added as the sole carbon source for the bacteria, are growth-linked biodegradation tests, thus, it is expected that there will be an increase in bacterial abundance in the test bottles with aniline degradation (Alexander, 1981).

4. Recommendations and conclusion

The purpose of this study was to develop a recommendation of either aging or not aging seawater before testing, and whether using deep water or surface water to give reproducible and high biodegradation results for the positive reference control aniline. However, the results were contradictory, only supporting the pool of evidence of the variability of biodegradation capacities found in microbial communities. Introducing a complex microbial community from natural seawater to laboratory conditions affects the microbial dynamics in a way that cannot be predicted, and that can affect the biodegradation potential of this community.

Even performing the same biodegradation test, at the same lab with the same operator with seawater from the same source and water from different depths at the same day, gives variation in biodegradation results. The microbial community changed over time both in storage cans and test bottles. This highlights the need for a more robust and reproducible microbial inoculum for these tests, consisting of a diverse community with high resilience to change.

Author contributions statement

Aina Charlotte Wennberg: Conceptualization; Investigation; Formal analysis; Writing – original draft. **Sondre Meland**: Formal analysis; Writing – review & editing. **Merete Grung**: Writing – review & editing. **Adam Lillicrap**: Conceptualization; Supervision; Writing – review & editing.

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.

Acknowledgments

This work was funded by Research Council of Norway, contract number 160016. Grphical abstract was created with BioRender.com.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemosphere.2022.134476.

References

Alexander, M., 1981. Biodegradation of chemicals of environmental concern. Science 211 (4478), 132–138.
Buysschaert, B., Kerckhof, F.M., Vandamme, P., De Baets, B., Boon, N., 2018. Flow cytometric fingerprinting for microbial strain discrimination and physiological characterization. Cytometry 93A (2), 201–212.
Cousins, I.T., Ng, C.A., Wang, Z., Scheringer, M., 2019. Why is high persistence alone a major cause of concern? Environ. Sci.: Processes & Impacts 21 (5), 781–792.
Davenport, R., Curtis-Jackson, P., Dalkmann, P., Davies, J., Fenner, K., Hand, L., McDonough, K., Ott, A., Ortega-Calvo, J.J., Parsons, J.R., Schaffer, A., Sweetlove, C., Trapp, S., Wang, N., Redman, A., 2022. Scientific concepts and methods for moving persistence assessments into the 21(st) Century. Integrated Environ. Assess. Manag. 1–34. https://doi.org/10.1002/ieam.4575. Online ahead of print. In press.
ECETOC, 2003a. E.C.f.E.a.T.o. In: Persistence of Chemicals in the Environment. Technical Report No. 90. . Chemicals. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium, p. 158.
ECETOC, 2003b. E.C.f.E.a.T.o. In: Persistence of Chemicals in the Environment. Technical Report No. 90. Chemicals. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium, p. 158.
Ecetoc, 2017. Workshop Report no.34 – Improvement of the OECD 306 Screening Test, p. 28.
ECHA, 2017. Guidance on Information Requirements and Chemical Safety Assessment. Chapter R.7b: Endpoint specific guidance, Helsinki, Finland.
Ellis, B., Haaland, P., Hahne, F., Meur, N.L., Gopalakrishnan, N., Spidlen, J., Jiang, M., Finak, G., 2020. flowCore: flowCore: Basic Structures for Flow Cytometry Data. R package, version 2.0.1.
Fuhrman, J.A., Cram, J.A., Needham, D.M., 2015. Marine microbial community dynamics and their ecological interpretation. Nat. Rev. Microbiol. 13 (3), 133–146.
Guenther, S., Koch, C., Huebschmann, T., Roeske, I., Mueller, R.A., Bley, T., Harms, H., Mueller, S., 2012. Correlation of community dynamics and process parameters as a tool for the prediction of the stability of wastewater treatment. Environ. Sci. Technol. 46 (1), 84–92.
Hammes, F., Berney, M., Wang, Y., Vital, M., Koester, O., Egli, T., 2008. Flow-cytometric total bacterial cell counts as a descriptive microbiological parameter for drinking water treatment processes. Water Res. 42 (1–2), 269–277.
Kinet, R., Dzaomuho, P., Baert, J., Taminiau, B., Daube, G., Nezer, C., Brostaux, Y., Nguyen, F., Dumont, G., Thonart, P., Delvigne, F., 2016. Flow cytometry community fingerprinting and amplicon sequencing for the assessment of landfill leachate cellulolytic bioaugmentation. Bioresour. Technol. 214, 450–459.
Koch, C., Guenther, S., Desta, A.F., Huebschmann, T., Mueller, S., 2013. Cytometric fingerprinting for analyzing microbial intracommunity structure variation and identifying subcommunity function. Nat. Protoc. 8 (1), 190–202.
Koch, C., Harnisch, F., Schroder, U., Muller, S., 2014. Cytometric fingerprints: evaluation of new tools for analyzing microbial community dynamics. Front. Microbiol. 5.
Kowalczyk, A., Martin, T.J., Price, O.R., Snape, J.R., van Egmond, R.A., Finnegan, C.J., Schafer, H., Davenport, R.J., Bending, G.D., 2015. Refinement of biodegradation tests methodologies and the proposed utility of new microbial ecology techniques. Ecotoxicol. Environ. Saf. 111, 9–22.
Ladau, J., Sharpton, T.J., Finucane, M.M., Jospin, G., Kembel, S.W., O'Dwyer, J., Koeppel, A.F., Green, J.L., Pollard, K.S., 2013. Global marine bacterial diversity peaks at high latitudes in winter. ISME J. 7 (9), 1669–1677.
Lambrecht, J., Schattenberg, F., Harms, H., Mueller, S., 2018. Characterizing Microbiome Dynamics - Flow Cytometry Based Workflows from Pure Cultures to Natural Communities, 137. Jove-Journal of Visualized Experiments.
Liu, Z., Cichocki, N., Hübschmann, T., Süring, C., Ofiţeru, I.D., Sloan, W.T., Grimm, V., Müller, S., 2019. Neutral mechanisms and niche differentiation in steady-state insular microbial communities revealed by single cell analysis. Environ. Microbiol. 21 (1), 164–181.
Liu, Z.S., Cichocki, N., Bonk, F., Gunther, S., Schattenberg, F., Harms, H., Centler, F., Muller, S., 2018. Ecological stability properties of microbial communities assessed by flow cytometry. mSphere 3 (1).
Martin, T.J., Snape, J.R., Bartram, A., Robson, A., Acharya, K., Davenport, R.J., 2017. Environmentally relevant inocula concentrations improve the reliability of persistent assessments in biodegradation screening tests. Environ. Sci. Technol. 51 (5), 3065–3073.
Nyholm, N., Kristensen, P., 1992. Screening methods for assessment of biodegradability of chemicals in seawater - results from a ring test. Ecotoxicol. Environ. Saf. 23 (2), 161–172.
OECD, 1992a. Test No. 301: Ready Biodegradability.
OECD, 1992b. Test No. 306: Biodegradability in Seawater.
Ospar, 2012. commission, O. In: OSPAR Guidelines for Completing the Harmonised Offshore Chemical Notification Format (HOCNF) (OSPAR Agreement: 2012-05. Update 2021), ed.
Ospar, 2015. In: OSPAR Guidelines for Completing the Harmonised Offshore Chemical Notification Format (HOCNF), ed. OSPAR.
Ott, A., Martin, T.J., Acharya, K., Lyon, D.Y.-M., Robinson, N., Rowles, B., Snape, J.R., Still, I., Whale, G.F., Albright III, V.C., Bäverbäck, P.B., Best, N., Commander, R., Eickhoff, C., Finn, S., Hidding, B., Maischak, H., Sowders, K.A., Taruki, M., Walton, H.E., Wennberg, A.C., Davenport, R.J., 2020a. Multi-laboratory validation of a new marine biodegradation screening test for chemical persistence assessment. Environ. Sci. Technol.
Ott, A., Martin, T.J., Snape, J.R., Davenport, R.J., 2020b. Increased cell numbers improve marine biodegradation tests for persistence assessment. Sci. Total Environ. 706.

Ott, A., Martin, T.J., Whale, G.F., Snape, J.R., Rowles, B., Galay-Burgos, M., Davenport, R.J., 2019. Improving the biodegradability in seawater test (OECD 306). Sci. Total Environ. 666, 399–404.

Prest, E.I., El-Chakhtoura, J., Hammes, F., Saikaly, P.E., van Loosdrecht, M.C.M., Vrouwenvelder, J.S., 2014. Combining flow cytometry and 16S rRNA gene pyrosequencing: a promising approach for drinking water monitoring and characterization. Water Res. 63, 179–189.

Prest, E.I., Hammes, F., Kotzsch, S., van Loosdrecht, M.C.M., Vrouwenvelder, J.S., 2013. Monitoring microbiological changes in drinking water systems using a fast and reproducible flow cytometric method. Water Res. 47 (19), 7131–7142.

Props, R., Monsieurs, P., Mysara, M., Clement, L., Boon, N., 2016. Measuring the biodiversity of microbial communities by flow cytometry. Methods in Ecology and Evolution 7 (11), 1376–1385.

Quiroga, M.V., Mataloni, G., Wanderley, B.M.S., Amado, A.M., Unrein, F., 2017. Bacterioplankton morphotypes structure and cytometric fingerprint rely on environmental conditions in a subAntarctic peatland. Hydrobiologia 787 (1), 255–268.

Rogers, W.T., Holyst, H.A., 2009. FlowFP: a bioconductor package for fingerprinting flow cytometric data. Advances in bioinformatics, 2009.

Sjoholm, K.K., Dechesne, A., Lyon, D., Saunders, D.M.V., Birch, H., Mayer, P., 2022. Linking biodegradation kinetics, microbial composition and test temperature - testing 40 petroleum hydrocarbons using inocula collected in winter and summer. Environ. Sci.-Process Impacts 24 (1), 152–160.

ter Braak, C.J., Smilauer, P., 2018. Canoco Reference Manual and User's Guide: Software for Ordination. Microcomputer Power, Ithaca, NY, USA, version 5.12.

Thouand, G., Capdeville, B., Block, J.C., 1996. Preadapted inocula for limiting the risk of errors in biodegradability tests. Ecotoxicol. Environ. Saf. 33 (3), 261–267.

Van Nevel, S., Buysschaert, B., De Gusseme, B., Boon, N., 2016. Flow cytometric examination of bacterial growth in a local drinking water network. Water Environ. J. 30 (1–2), 167–176.

Van Nevel, S., Buysschaert, B., De Roy, K., De Gusseme, B., Clement, L., Boon, N., 2017. Flow cytometry for immediate follow-up of drinking water networks after maintenance. Water Res. 111, 66–73.