Cefas audits of HOCNF Substance Test Reports: Review and Recommendations

Executive Summary

Cefas registers Offshore Chemicals for use in The Netherlands, based on the laboratory data provided by the chemical supply industry using the Harmonised Offshore Chemical Notification Format (HOCNF). Audits of the original laboratory reports underpinning the registrations of selected products have been conducted annually on behalf of the State Supervision of Mines (SSM) and the Netherlands Oil and Gas Exploration & Production Association (NOGEPA) since 2007. Although the reports detail nominally recognised (typically OECD) studies conducted in accordance with Good Laboratory Practice, the audits identified various unsatisfactory features in the reports, which comprise studies covering bioaccumulation potential, biodegradability and toxicity to marine species. This report discusses the issues commonly identified with each type of test and includes recommendations for the benefit of both test laboratories and chemical suppliers.

Contents

1.0 Introduction

Cefas have conducted audits on the Harmonised Offshore Chemical Notification Format (HOCNF) and their underlying reports from 2007 until 2017 inclusive. These audits focussed on 10 products each year and were based on the criteria originally devised in 2006 by TNO for State Supervision of Mines (SSM). These audits have been the subject of Netherlands papers to the OSPAR Offshore Industry Committee in 2008 and 2009.

Successive audits have recorded failings in the quality of test reports, both minor and major, in a number of areas. In serious cases, these have prompted Cefas to request that tests should be repeated. However, after discussions with SSM and NOGEPA, it was agreed that increased feedback to the chemical supply industry and its associated test laboratories was required in order to drive improvements in this area. This report has been prepared for that purpose and is intended to be circulated to chemical suppliers and their test laboratories.

The report includes four main sections. These comprise summaries of the key issues identified in tests associated with persistence, bioaccumulation and toxicity respectively, which are preceded by a section devoted to issues of a more general nature. In addition, appendices are included which comprise a series of tables itemising the checks carried out by Cefas in each audit. These describe the checks associated with the conduct of i) the test laboratory and ii) the supplier responsible for completing the HOCNF form respectively.

2.0 General issues with Test Reports

2.1 Test substance specification

The EU principles of Good Laboratory Practice (GLP) 2004 state in section 6 on Test and reference items:

> *6.2 Characterisation*
>
> *1. Each test and reference item should be appropriately identified (e.g. code, chemical abstracts service registry number (CAS number), name, biological parameters).*
>
> *2. For each study, the identity, including batch number, purity, composition, concentrations, or other characteristics to appropriately define each batch of the test or reference item should be known.*
>
> *3. In cases where the test item is supplied by the sponsor, there should be a mechanism, developed in cooperation between the sponsor and the test facility, to verify the identity of the test item subject to the study.*
>
> *4. The stability of the test and reference items under storage and test conditions should be known for all studies.*
>
> *5. If the test item is administered or applied in a vehicle, the homogeneity, concentration and stability of the test item in that vehicle should be determined. For test items used in field studies (e.g. tank mixes), these may be determined through separate laboratory experiments.*
>
> *6. A sample for analytical purposes from each batch of test item should be retained for all studies except short-term studies.*

The interpretation of this text that has been employed in most GLP reports that have been audited is one in which paragraph 3 is used in isolation. This paragraph appears to be considered by the test laboratories to be an alternative to paragraphs 1 and 2. In the resulting reports, the test laboratory has taken no responsibility for the characterisation of the test substance but identified the study sponsor as having that responsibility. CAS numbers and chemical names are not always included, with the test substance being identified by a code or trade name instead.

To have a high level of confidence that the study is adequate to fulfil the requirements for a data point on the HOCNF, it is important to have a full and clear understanding of what the test item was, and that the test item named in the report was indeed the one that was tested. As such the requirements in paragraphs 2 and 3 must always be met, whilst SSM also require Cefas to seek evidence that the study director had confirmed the identity of the test item received from the study sponsor or a third party. Very few reports were found to meet these requirements.

In order to take account of the scale of non-conformance in this area, Cefas has agreed the following future acceptance criteria with SSM:

a) New Test Reports (i.e. those commissioned after January 1st 2019)

Test reports must comply with all aspects of section 6.2 of the GLP Regulations with respect to Test Substance Characterisation and must include evidence that the study director had confirmed the identity of the test item received from the study sponsor or a third party. It should be stressed that the studydirector has the responsibility to receive all the information about the substance identification from the supplier. In case the information provided is insufficient, the studydirector can reject the sample or can make a statement in the testreport which information is missing. The authorities can reject the report if information on the substance identification is missing.

b) Existing Test Reports (i.e. those commissioned prior to January 1st, 2019)

When Cefas assesses a test report and finds it to be deficient in terms of test substance characterisation, the supplier would be required to submit details of the test substance in accordance with the ECHA guidance (https://www.echa.europa.eu/documents/10162/23036412/substance_id_en.pdf/ee696bad-49f6-4fec-b8b7-2c3706113c7d). If a supplier is unable to provide the required information, to characterise the test substance, Cefas, will make a decision on a case by case basis if it accepts the reportwhereby the decision to accept/reject a test report, will depend upon factors such as the extent of the information deficiencies

2.2 Deviations from Test Protocol

It is not uncommon for Laboratories to conduct tests according to their own interpretations of OECD Protocols, instead of employing strict adherence to those Protocols as written. Laboratories should note that if deviations from the protocol occur, the report should detail those deviations and include a comment, with justification, on their likely effect upon the results of the study.

2.3 Reference to Internal Standard Operating Procedures

Certain Laboratories make frequent reference in their reports to particular procedures having been conducted according to internal Standard Operating Procedures (SOPs) for that Laboratory. However, unless those SOPs are available via the Laboratory's web site, this approach effectively renders parts of the study inaccessible to external bodies When conducting a test according to a SOP, this should be made available to the authorities upon request.

3.0 Bioaccumulation

3.1 Introduction

The HMCS allows information on the bioaccumulation potential to be provided by suppliers in the form of a bioconcentration factor (BCF) generated in fish (OECD 305) or bivalves (ASTM E1022). Alternatively, bioaccumulation potential may be estimated from the log P_{ow} determined by HPLC (OECD 117) or shake flask (OECD 107) methods, or by suitably validated calculation methods, i.e. quantitative structural activity relationships (QSAR).

For the most part, the data submitted on the HOCNF to fill the bioaccumulation end point are OECD 117 or QSAR data and therefore the remainder of this section concentrates on these study types.

3.2 Methods of detection

The OECD 117 Protocol advocates the use of a UV Detector tuned to a wavelength of 210nm or a Refractive Index Detector, on the basis that they are "applicable to the wide variety of chemical groups". However, it must be appreciated that neither of these detection modes offers a universal response, and as a result, the percentages of the total peak area represented by each peak may not reflect the actual composition of the sample. This is particularly relevant to the industrial grade substances used in offshore chemicals since the substances may comprise multiple compounds, whose different detector responses ensure that the principal component may appear as only a minor peak in the chromatogram. This problem becomes particularly acute when the UV Detector is used for compounds which lack chromophoric groups, since they will be undetectable, and the results (if any) will be unreliable. Non-detection of such compounds is a lesser concern for the RI detector, since its response is driven by the refractive index of the mobile phase as it leaves the column. However, it suffers from the insensitivity that is common to all bulk property detection methods. Additional confidence can be gained where the two detectors are linked in series, but this approach cannot be relied upon if components of the sample are strongly retained on the column and not eluted. Such circumstance may result in the Log P_{ow} being derived from a fast-eluting impurity and the value obtained being totally unrepresentative of the bulk of the test substance.

As a result of the foregoing issues, it is essential that the test laboratory is aware of and understands the nature of the chemical it is testing. Where evidence of such understanding is absent, it is often difficult for the regulator to have any level of confidence that the Log P_{ow} derived is valid for the test substance.

3.3 Standards

The OECD 117 protocol (13 April 2004, page 3, item 14) states that:

> *In order to correlate the measured capacity factor k of a compound with its P_{ow}, a calibration graph using at least 6 points has to be established. It is up to the user to select the appropriate reference compounds. It is preferable that these should be structurally related to the test substance. Whenever*

possible, at least one reference compound should have a P_{ow} above that of the test substance, and another a P_{ow} below that of the test substance. For log P_{ow} values below 4, the calibration can be based on data obtained by the Shake Flask method. For log P_{ow} values above 4, the calibration can be based on literature values if they correspond to calculated values.

In general, this is an aspect of the studies which is often disregarded by the test laboratories with many appearing to use the same standard set regardless of the test substance. This is again an item which may lead to questions over the reliability of the study particularly where the result of the study is extrapolated rather than interpolated from the curve.

3.4 QSAR validation

The OECD principles of QSAR state that.

To facilitate the consolidation of a (Q)SAR model for regulatory purposes, it should be associated with the following information:

1. *A defined end point*
2. *An unambiguous algorithm*
3. *A defined domain of applicability*
4. *Appropriate measures of goodness-of-fit, robustness and predictivity*
5. *A mechanistic interpretation, if possible*

As generally speaking the software used for the QSAR calculations is downloaded from the web from regulatory organisations (EPISuite from the EPA, OECD QSAR toolbox from the OECD), items 1, 2, and 4 are generally well accounted for within the model. However, item 3, which refers to the validation of the applicability domain of the QSAR calculation, is generally not conducted, nor is item 5. There appears to be a general ignorance of these requirements within test laboratories and chemical suppliers (from both large multinational companies and SMEs).

Without a validation of the applicability domain of the QSAR in respect of the test substance there is no way of knowing whether the QSAR algorithm over- or under-estimates the Log P_{ow} and the extent to which the calculated value is likely to deviate from the true value.

A further complication here is the use of QSAR to predict BCF directly for difficult to test substances, such as surfactants. Again, there is a need for validation of the applicability domain here, but it is also important to understand whether the underlying algorithm is based on a log P_{ow} QSAR and if this is an appropriate method to predict the BCF of the difficult to test substance (i.e. whether hydrophobic or electrostatic interactions are driving the bioconcentration process). This requires a thorough understanding of the chemistries involved and exemplifies why the provision of a mechanistic interpretation is beneficial from a regulatory standpoint.

3.5 Recommendations: Bioaccumulation

- Laboratories employing the OECD 117 should:

- Consider whether the detection method used is appropriate for the test substance and that all components are fully accounted for in the results;
- Ensure that all of the peaks used in formulating the results are related to the test substance and are not artefacts of the method;
- Ensure that weighted average log P_{ow} values are only determined for homologous series.

- Laboratories employing QSAR data should:
  - Ensure the applicability domain of the QSAR algorithm has been fully validated for the test substance;
  - Ensure that the underlying algorithm is appropriate for the determination and fully complies with the OECD principles of QSAR validation.

4.0 Biodegradation

4.1 Introduction

The majority of studies submitted for the biodegradation endpoint are derived from the OECD 306 marine biodegradation test, with a few OECD 301 studies submitted. Other studies generally employ the Marine BODIS method, as described in Appendix 6 of the OSPAR HOCNF Guidelines.The issues described below are common to all types of study, but the following will concentrate on the OECD 306 protocol.

4.2 Method Selection

The OECD 306 method may be conducted as a “closed bottle” or “shake flask” test, with the choice of test influenced by the characteristics of the test substance. In the closed bottle test, oxygen demand is used as the analytical method for following the extent of degradation for the test substance. This methodology ideally should be based upon the calculated value for the oxygen demand of the test substance added (Theoretical Oxygen Demand ThOD). This calculation is based upon the elemental composition of the test substance. It is permitted (if the elemental composition of the test substance is unknown) for the chemical oxygen demand (COD) to be determined and used as the basis for the test. It is however possible for the oxidation of the test substance to be incomplete in the COD determination and for this to result in an over estimate of the biodegradation of the test substance. Cefas will reject reports if the ThOD has not been used without good reason. If the test substance is of Unknown or Variable Composition (UVCB) a ThOD can still be calculated instead of a COD being experimentally determined; however, the choice of chemical formula used in the calculations must be fully justified.

In the shake flask test, degradation is monitored by determining the level of dissolved organic carbon. It is important to consider these data carefully, particularly when low concentrations of test substance are being studied. Variations in the analytical method may result in fluctuating results when background values are subtracted from sample value and a trend in the biodegradation of the substance is therefore difficult to discern. This is particularly an issue when low levels of substance are used, and the precision of the raw data is limited, when the result can fluctuate between pass and fail values when compared to HMCS pre-screening criteria.

A further complication here is many test laboratories’ inability to account for a false positive result as an artefact of nitrification. This is also particularly an issue for COD-based studies. Irrespective of the method employed, the presence of solvents is a potential source of erroneous results.

4.3 Temperature of incubation

The Closed Bottle version of the OECD 306 test should be conducted at a temperature controlled to ±1 ^{o}C within the range 15 to 20^{o}C. Cefas have observed that some laboratories appear to believe that the requirement is 20 +/- 1 ^{o}C, with the result that a number of tests are conducted at the higher end of the permitted range, and sometimes exceed the upper limit. Where this occurs, it is imperative that the study report provides full details of the extent and duration of the deviation, its timing, and the anticipated effect upon the results. In the absence of such information, Cefas may conclude that the higher temperature will have resulted in an unrealistic level of biodegradation to have occurred and consider that the report is unreliable.

4.4 Variability of OECD 306 Marine biodegradation tests

The OECD 306 marine biodegradation test is well known for providing highly variable results. This is to a large extent due to the low biomass present in the test flasks used to the conduct the test; which is a result of

i) the low number of colony forming units (cfu) detected in seawater, which varies with season and location (10^3 to 10^8 cfu/ml in seawater *cf* 10^8 CFU/ml required to get a consistent result ECETOC 2001, 2007, 2009),
ii) the low volume of sea water in each flask (which reduces the probability of the inclusion of microbes capable of degrading the test substance) and
iii) this test's underlying requirement for degradation to be based upon Monod growth kinetics of a competent degrader of the test substance.

In addition, the detection methods used in this type of study often require environmentally unrealistic test substance concentrations to be studied. As such a surface water simulation test conducted to the OECD 309 protocol would often be a more appropriate test provided metabolites were considered in the analysis of the test samples.

4.5 Recommendations: Biodegradation

Laboratories conducting biodegradability studies should pay close attention to the following:

- The suitability of the method chosen for the test substance.
- The nature of the test substance should also be carefully evaluated to ensure that the biodegradation seen is not primarily that of any carrier solvent present.
- The experimental conditions used should be carefully evaluated particularly in respect of the temperature of incubation.

5.0 Toxicity

5.1 Extraneous Factors

Many factors may affect the results of a toxicity test with marine organisms. These factors may be due to the characteristics of the water, test species or experimental design. It is for these reasons that standardised test protocols are used in regulatory toxicology. A fundamental part of all these marine aquatic toxicity tests is the requirement to document test water quality at regular intervals. For natural water the parameters analysed should include the salinity, pH, dissolved oxygen, ammonia nitrogen, nitrite, pesticides and metals. In the case of artificial sea water, it is important to conduct a full analysis on the laboratory water supply used for the toxicity tests as well as using chemicals of a suitable grade to prepare the artificial sea water.

Many laboratories do not maintain stocks of the various test species in-house and buy these in shortly before commencing the test. It is important to allow these organisms an appropriate amount of time to recover from the stress of transportation and to acclimatise to their new environment. Fully reviewing the health and fitness of the organisms is important before commencing testing. For many studies these organisms must be in a particular life stage or be a particular size and therefore there must be clear evidence of this in the final report to give confidence in the reliability of the data.

This fundamental information is not always available in the test report for a particular toxicity test and in many cases where it is present it is historic information or a general statement.

5.2 Statistics

Regulatory toxicity tests are designed using statistics to ensure that the results that are generated are adequate and reliable. This design is inherent in the number of organisms used, the treatment groups and the order of treatment with test substance. The data generated from the test must therefore be analysed statistically to derive the median lethality or effects concentration as well as the no observed effects concentration. The requirement to describe the statistical methods used for the data analysis as well as a full and clear presentation and discussion of the results is vitally important to the interpretation of the toxicity data.

5.3 QSARs

Increasing use is being made of QSARs for ecotoxicity endpoints, replacing laboratory measurements. Where such methods are employed, they should be conducted in accordance with OECD principles (see Section 3.4).

5.4 Recommendations: Toxicity

Laboratories conducting toxicity measurements should:

- Focus on ensuring a full set of basic information is available with respect to the test species, water quality, handling of the test species and study setup and conduct to ensure that the study is relevant and reliable.
- Ensure that correct statistical methods have been followed and correctly applied to the raw data to ensure that the final results are reliable and adequate to fill the data point.

6.0 General recommendations

Recommendations regarding the conduct of specific tests types have been included in the preceding sections and should be followed in the future in order to assist regulatory acceptance of test results obtained from such methods. In addition, Cefas recommends that on a broader level, a closer collaboration between the chemical supplier (i.e. test sponsor) and the test laboratory should be sought, and particularly:

i. Suppliers should provide testlaboratories with information about product identity as per GLP requirements. If this information is lacking, authorities can reject the testreport.
ii. Laboratories should pay closer attention to the exact requirements of the protocols that they claim to be following and discuss any deviations in the test reports.
iii. Suppliers should recognise that test protocols generally state clearly what information should be expected to be included in the test report and be prepared to challenge any laboratory that prodices test reports that lack such information.
iv. Suppliers should also appreciate that for certain tests that are required to be conducted to facilitate an HMCS registration (especially OECD 117) the laboratory will require understanding of the test substance in order to carry out a test that is fit for purpose.

7.0 References

Directive 2004/10/EC of the European Parliament and council of 11 February 2004 on the harmonisation of laws, regulations and administrative provisions relating to the application of the principles of good laboratory practice and the verification of their applications to tests on chemical substances (codified version)

Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a

European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC

OECD Principles for the validation, for regulatory purposes, of (quantitative) structure-activity relationship models 37th joint meeting of the chemical committee and working party on Chemicals, pesticides and biotechnology. November 2004

OSPAR Guidelines for Toxicity Testing of Substances and Preparations Used and Discharged Offshore (Reference number: 2005-12)

OSPAR Decision 2000/2 on a Harmonised Mandatory Control System for the Use and Reduction of the Discharge of Offshore Chemicals, as amended by OSPAR Decision 2005/1.

OSPAR Recommendation 2000/4 on a Harmonised Pre-screening Scheme for Offshore Chemicals as amended by OSPAR Recommendations 2008/1, 2010/04; 2016/04 and 2017/01

OSPAR Recommendation 2000/5 on a Harmonised Offshore Chemical Notification Format (HOCNF) as amended by OSPAR recommendations 2005/3, 2008/2, 2010/03 and 2014/17.

ECETOC TR 082: Risk Assessment in Marine Environments December 2001

ECETOC WR 10: Workshop on Biodegradation and Persistence**.** September 2007

ECETOC TR 108: Collation of Existing Marine Biodegradation Data and its Use in Environmental Risk

assessment December 2009

ECETOC WR 34: Improvement of the OECD 306 screening test September 2017

OECD Guidelines for Testing of Chemicals, 1995-107. Partition Coefficient (n-octanol/water): Shake Flask

Method

OECD Guidelines for Testing of Chemicals, 2004 – 117 Partition Coefficient (n-octanol/water), High Performance Liquid Chromatography (HPLC) Method

OECD Guidelines for Testing of Chemicals, 1992– 301, Ready Biodegradability

OECD Guidelines for Testing of Chemicals, 1992 – 306, Biodegradability in Seawater

OECD Guidelines for Testing of Chemicals, 1996 – 305 Bioconcentration: Flow-through Fish Test

ASTM E1022-94 (2007) Standard guide for conducting Bioconcentration Tests with Fishes and Saltwater

Bivalve mollusks

ISO/DIS 10253 Water Quality- Marine algal growth inhibition test with S*keletonema costatum* and

Phaeodactylum tricornutum

ISO 14669:1999 Water Quality-Determination of acute lethality to marine copepods (copepoda, Crustacean)

Part B of the OSPAR Protocols on Methods for *Cyprinodon variegatus (juveniles)* the Testing of Chemicals Used in the Offshore Industry (published by OSPAR in 1995, available from the OSPAR web site www.ospar.org)

Part A of the OSPAR Protocols on Methods for the Testing of Chemicals Used in the Offshore Industry A sediment bioassay using an *Amphipod Corophium sp*

KLIMISCH HJ, ANDREAE M, TILLMANN U, REGULATORY TOXICOLOGY AND PHARMACOLOGY 25(1), 1–5 (1997) A Systematic Approach for Evaluating the Quality of Experimental Toxicological and Ecotoxicological Data 1

ECHA, 2008: Guidance on information requirements and chemical safety assessment Chapter R.4: Evaluation of available information

APPENDIXES: Cefas Criteria for Netherlands Report Auditing

Outline

The objectives of the audit are to

1. Determine that values have been accurately transcribed from the final reports to the current HOCNF on file at Cefas.
2. Determine whether the reports submitted are conducted to GLP and meet the REACH requirements for reliability as specified in the Guidance for Information Requirements Document R4.
3. Determine whether the reports are adequate for the purposes of filling the HOCNF data points for which they are submitted.

Additional detail of specific criteria that are addressed in the course of performing these checks is provided in the two following sections.

APPENDIX 1: Supplier Criteria

The following sections detail the checks that are performed by Cefas in order to ensure that the chemical supplier has commissioned appropriate tests and that the results from these tests have been transcribed accurately onto the HOCNF form.

A1.1 Adequacy and Relevance Check

Item	Requirement	Comment	Concern level
1	Is the report conducted to an internationally accepted standardised test protocol that meets the requirements of the data point with respect to pre-screening and CHARM requirements?	As defined in HOCNF guidance	Yes: No concern No: Fail test
2	Or are the methods used to derive the data point scientifically justified and robust and provide data that meets the requirements of the data point with respect to pre-screening and CHARM requirements?	Do studies meet HOCNF / CHARM requirements with appropriate methods not described in HOCNF guidance?	Yes: No Concern No: Fail test

A1.2 Transcription check

Item	Requirement	Comment	Concern level
1	Is the report uniquely and correctly identified in the HOCNF?	Care should be taken in assessing this point as reports often have a range of unique identifier numbers on them that are confusing and pragmatism is therefore needed in making this assessment	No: New HOCNF required Yes: Continue audit
2	Have the data been selectively transcribed to the report?	If there is a concern that selective reporting has occurred to influence the outcome of an assessment this should immediately be brought to the attention of SSM. In particular, care should be applied to the handling of toxicity data for substances supplied as aqueous solutions. Suppliers must indicate the concentration of such solutions to ensure that they are correctly assessed.	No: Continue audit Yes: New HOCNF required
3	Have all values that are required on the HOCNF been correctly transcribed from the report?	This test is intended to look at the quality systems used in producing the HOCNF	No: New HOCNF required Yes: Continue audit
4	Has the confirmation statement been signed in section 3?	The HOCNF is not complete	No: New HOCNF required Yes; No concern

A1.3 Test Substance Characterisation Check

Item	Requirement	Comment	Concern level
1	Test substance fully identified with chemical name, CAS*, EC number*, purity and other relevant descriptors	*CAS and EC number should be provided where applicable. If this information is not available in the report it must be provided by the chemical supplier. Allowance must be made for the complexity of the test substance when considering identification requirements. Characterisation details should be in accordance with ECHA *Guidance for identification and naming of substances under REACH and CLP*.	No: Report is potentially unreliable: request information from supplier Yes: Continue audit

A1.4 GLP Compliance Check

Item	Requirement	Comment	Concern level
1	Study directors signature on GLP compliance statement	If absent this is not a final report	No; request that a Klimisch assessment is carried out. Yes: Continue audit
2	QA signature on audit report	If absent this is not a final report	No: request that a Klimisch assessment is carried out. Yes: Continue audit

A1.5 Reliability Evidence Check

Item	Requirement	Comment	Concern level
1	Statement of Klimisch assessment is requested and sent	The reliability of a report is assessed after being called into question for not being GLP compliant.	Yes: Continue audit No: report unreliable
2	Evidence is available that the person completing the Klimisch assessment is appropriate qualified, or trained or experienced.	This is a REACH requirement	Yes; continue audit No: assessment and report are potentially unreliable
3	Statement fully describes the deficiencies of the work and awards Klimisch score appropriately	Information on the Klimisch scoring system can be found on the ECHA website.	Yes: Continue audit No: report is potentially unreliable
4	Klimisch assessment results in a score of 1 or 2		Yes: Continue audit No: report unreliable

NOTE: Under REACH, new ecotoxicological test data must be conducted to GLP standards, but acceptance of existing tests is permitted where these are adequately reliable. The Reliability Evidence Check is conducted during the Audit whenever a test report is encountered that is not GLP-compliant.

The current HOCNF requires the signatory to confirm that:

> "the laboratory test results and data that form the basis of this document are either in compliance with the requirements of the relevant REACH registration, or in compliance with the European Chemicals Agency (ECHA) 'Guidance on information requirements and Chemical Safety Assessment', Chapter R4: Evaluation of available information, May 2008 (as amended)".

This check is intended to test whether the supplier has evaluated their data for compliance with this statement or not. i.e. it is the suppliers' responsibility to conduct this assessment and as such this test is looking for proof that this has been done.

APPENDIX 2: Study specific audit requirements

The following sections detail the checks that are performed by Cefas in order to ensure that test laboratories have conducted the relevant protocols correctly.

All tests in the following sections should be conducted to make a decision about the reliability of each report. This section of the guidance is designed to standardise the Klimisch assessment of all reports that have passed the initial checks. The purpose of these checks is to determine if there are any deficiencies in the methodologies used and to determine how these deficiencies impact the reliability, adequacy and relevance of the study for the HMCS data requirements.

It is noted that a failure to meet test requirements can lead to the report being deemed to be unreliable. Where this occurs, the nature of the non-conformance(s) may result in an endpoint that is indicative of:

a) A greater hazard than that obtained from a correctly-performed test
b) A lesser hazard than that obtained from a correctly-performed test
c) A hazard that may be greater or less than that obtained from a correctly-performed test

It is important to note that a report that is deemed to be unreliable may still be considered adequate for regulatory purposes, if it indicates a greater hazard.

A2.1 OECD 117

Item	Requirement	Comment	Concern level
2	Evidence that all components of the substance are fully considered	Consideration should be given to possible non-elution of compounds in the test substance, in view of their polarity and the strength of the eluent.	No. Report is unreliable
3	Evidence that the standards used are appropriate for the test substance	Refer to OECD 117 protocol. Results should be interpolated from the curve	No: report is potentially un reliable
4	All values determined are between 0 and 6	Values greater than 6 may be used for pre-screening purposes but should not be used for CHARM calculations	No: Method is not adequate or relevant for the determination of Log Pow but values of greater than 6 may be accepted based on the judgement of the senior scientist
5	Method is suitable for the test substance	Substance is inorganic, a strong acid or base, metal complex, or surfactant (as defined by OSPAR)	No: Method is not adequate or relevant
6	Substance reacts with the eluent		Yes: Method is not adequate
7	Duplicate measurements made and fall within 0.1 log units	This is a requirement to demonstrate the repeatability of the method.	No: Method is not adequate
8	Mobile phase meets the requirements of OECD 117	Refer to OECD 117 protocol.	No: Method is not adequate
9	Dead time from column determined	Refer to OECD 117 protocol. It is important that the test substance is retained by the column for this method to produce reliable results	No: Method is not reliable
10	Equipment and method of use fully described (column, guard column, mode of operation, mobile phase, injection volumes, temperature, pH)	It is a requirement of GLP to clearly describe how a study was conducted. The amount of information presented must be carefully considered to determine whether the study is adequate	No: the report is not reliable
11	Method of calculation fully documented, traceable and correct	It should be possible to fully reconstruct how the final value was determined from the final report	No: the report is not reliable
12	All validity criteria met	Refer to OECD 117 protocol.	No: the report is not adequate or reliable
13	Laboratory method employed	The OECD 117 Protocol includes an Annex on calculation methods. These should be treated as QSARs.	No: refer to Section 3.4.

A2.2 OECD 107

Item	Requirement	Comment	Concern level
1	Evidence that the detection method used is appropriate for the test substance	The report should clearly demonstrate that the detection method(s) used produce a reproducible and positive response to the components in the substance	No: Report is unreliable
2	Method is suitable for the test substance	The presence of impurities in the test substance can compromise the test result, especially if they have a high response factor with the detection method.	No. Report is unreliable
3	Full description of method is provided including centrifuge time and speed	It is a requirement of GLP to clearly describe how a study was conducted	No: The amount of information presented must be carefully considered to determine whether the study is adequate
4	Temperature is in the range 20 to 25°C and does not vary by more than ±1°C	Refer to OECD 107 protocol.	No: Study is not reliable, however the results may provide sufficient information to judge the substance to be bioaccumulative
5	all information relevant for the interpretation of the results, especially with regard to impurities and physical state of the substance	Refer to OECD 107 protocol.	No: the report is not reliable
6	pH of the water used and of the aqueous phase during the experiment; and where applicable justification for the use of buffers; composition, concentration and pH of the buffers; pH of the aqueous phase before and after the experiment;	Refer to OECD 107 protocol.	No: the report is not reliable

Item	Requirement	Comment	Concern level
7	Concentration data: the concentrations measured in each run (a total of 12 concentrations); - P_{OW} values and their mean for each set of test conditions and the overall mean (if there is the suggestion of concentration dependence of the partition coefficient, this should be noted); - the standard deviation of individual P_{OW} values about their mean; - the overall mean expressed as its logarithm to base 10; - the theoretical P_{OW} when it has been calculated or when the measured value is above 10^4.	The accuracy of the results is impaired by the use of high concentrations of the test substance. Further information is provided in the OECD 107 protocol.	No: the report is not reliable

A2.3 QSAR

Item	Requirement	Comment	Concern level
1	The SMILES notation is correct	This is the method by which the substance's structure is entered into the software	No: Report is not relevant
2	All components in the substance are accounted for	All components in a substance should be accounted for in line with the HOCNF guidance on composition	No: report is not adequate
3	The QSAR has a defined end point	For standardised regulatory models this is not an issue. Refer to OECD guidance for further information	No; QSAR is not adequate
4	The QSAR has an unambiguous algorithm	For standardised regulatory models this is not an issue. Refer to OECD guidance for further information	No; QSAR is not adequate
5	The QSAR applicability domain has been validated for the test substance	It is essential that the applicability of the QSAR for the substance is demonstrated to be at least conservative in the predictions that it makes.	No QSAR is not reliable
6	There are appropriate statistics for the QSAR goodness of fit, robustness and predictivity	For standardised regulatory models this is not an issue. Refer to OECD guidance for further information	No; QSAR is not adequate
7	The calculation method is appropriate for the substance	Check that the underlying method of calculation is appropriate for the test substance type e.g. a BCF QSAR based on log Pow prediction may not be adequate for a surfactant unless a clear relationship between BCF and log Pow has been demonstrated experimentally.	No: QSAR is not adequate

A2.4 OECD 306 or OSPAR Marine BODIS test protocols

Item	Requirement	Comment	Concern level
1	Appropriate reference standard used	Refer to OECD 306/ Marine BODIS test protocol.	No: the report is not reliable
2	Degradation of reference standard was as expected from ring test	Refer to OECD 306/ Marine BODIS test protocol.	No: the report is not reliable
	Seawater data are present: -collection location - date of collection; - depth of collection; - appearance of sample - turbid, etc.; - temperature at the time of collection; - salinity; - DOC; - delay between collection and use in the test - Heterotrophic microbes count - dissolved nitrate - dissolved ammonium - dissolved phosphate	It should be possible to fully reconstruct how the study was conducted. The parameters listed here are specified in the protocol and are important for allowing a full interpretation of the study data. The absence of any data point may not in itself make the study inadequate or unreliable but each dataset must be considered on its own merits.	No: the report is potentially unreliable
4	Test medium Composition and use fully described	It is a requirement of GLP to clearly describe how a study was conducted.	No: The amount of information presented must be carefully considered to determine whether the study is adequate
5	Incubation and analysis of samples fully described and within required parameters for the test	Refer to OECD 306/ Marine BODIS test protocol. It is a requirement of GLP to clearly describe how a study was conducted.	No: The amount of information presented must be carefully considered to determine whether the study is adequate
6	Results and calculation fully presented and described	It should be possible to fully reconstruct how the final value was determined from the final report.	No: the report is not reliable
7	Test result valid	Refer to OECD 306/ Marine BODIS test protocol for validity criteria	No: the report is not reliable

Item	Requirement	Comment	Concern level
8	COD used rather than ThOD	COD based studies may overestimate the extent of biodegradation due to incomplete oxidation of the test substance in the COD test. Results should be treated with caution	Yes: Treat result cautiously and consider possibility that the report is not reliable
9	Nitrification where appropriate taken into account	Refer to OECD 301 protocol for guidance; Nitrification can be a serious interference in the biodegradation of some substances resulting in a false positive result	No: Treat result cautiously and consider possibility that the report is not reliable
10	Sensitivity of analytical method appropriate for determination of pre-screening end points	The sensitivity of the analytical method should be evaluated to determine how well it discriminates between pass fail at cut off values (i.e. pre-screening)	If sensitivity is poor, lower result should be considered for use in hazard and risk assessment
11	ThOD calculated for UVCB substance	The calculations must be adequately explained/justified so they can be replicated.	No: the report is not reliable

A2.5 OECD 301

Item	Requirement	Comment	Concern level
1	Appropriate reference standard used	Refer to OECD 301 protocol.	No: the report is not reliable
2	Degradation of reference standard was as expected from ring test	Refer to OECD 301 protocol.	No: the report is not reliable
3	Test conditions fully reported inoculum: nature and sampling site(s), concentration and any pre-conditioning treatment; - proportion and nature of industrial waste water in sewage, if known; - test duration and temperature; - in the case of poorly soluble test substances, methods of preparation of test solutions/suspensions; - test method applied; scientific reasons and explanation for any change of procedure.	It should be possible to fully reconstruct how study was conducted. The collection handling and characterisation of the sewage sludge are important for allowing a full interpretation of the study data.	No: the report is not reliable
4	Incubation and analysis of samples fully described and within required parameters for the test	Refer to OECD 301 protocol: It is a requirement of GLP to clearly describe how a study was conducted	No: The amount of information presented must be carefully considered to determine whether the study is adequate
5	Results and calculation fully presented and described: - data in tabular form; - any observed inhibition phenomena; - any observed abiotic degradation; - specific chemical analytical data, if available; - analytical data on intermediates, if available; - the graph of percentage degradation against time for the test and reference substances, the lag phase, degradation phase, the 10-d window and slope (see Annex I of 301 protocol for definitions); - percentage removal at plateau, at end of test, and/or after 10-d window	The criteria listed here are stated in the OECD 301 protocol and their relevance and importance should be considered for each study. It is difficult to assess a report on any criterion that finishes with the clause "if available" however this is taken to indicate that this information is additional data with respect to this study type rather than essential	No: The amount of information presented must be carefully considered to determine whether the study is adequate
6	Test result valid	Refer to OECD 306/ ring test protocol for validity criteria	No: the report is not reliable

Item	Requirement	Comment	Concern level
7	COD used rather than ThOD	COD based studies may overestimate the extent of biodegradation due to incomplete oxidation of the test substance in the COD test. Results should be treated with caution	Yes: Treat result cautiously and consider possibility that the report is not reliable
8	Nitrification where appropriate taken into account	Refer to OECD 301 protocol for guidance; Nitrification can be a serious interference in the biodegradation of some substances resulting in a false positive result	No: Treat result cautiously and consider possibility that the report is not reliable
9	Sensitivity of analytical method appropriate for determination of pre-screening end points	The sensitivity of the analytical method should be evaluated to determine how well it discriminates between pass fail at cut off values (i.e. pre-screening)	If sensitivity is poor lower result should be considered for use in hazard and risk assessment

A2.6 ISO10253 Algal test

Item	Requirement	Comment	Concern level
1	Test organism fully described: Species origin, Strain number, Method of cultivation	Refer to test guideline. At a minimum the species, source and cultivation method are required	No test is not adequate
2	Synthetic or natural seawater medium preparation documented and salinity was 25 to 35%.	Refer to test guideline	No: test is not adequate
3	Evidence that the seawater medium used was not contaminated with any substances ((especially copper)	This is a fundamental requirement for any aquatic toxicity test as contamination with metals or pesticides can be a confounding factor in a toxicity test.	No: The toxicity seen may be due to substances other than those
4	Nutrient stock solution preparation fully recorded	Refer to test guidelines	No: there may be concern that organisms were not properly maintained
5	If test is conducted on substances containing cations, the effect of EDTA in the medium must be fully accounted for	EDTA will chelate cations, therefore perturbing the balance of ions in the test medium. This may affect growth.	No: test is not reliable
6	Test details fully reported: Start date, Duration, Concentrations tested, Composition of medium,	Refer to test guidelines	No: test is not adequate
7	Incubation equipment, methods, temperature (19-21°C), Light quality and intensity (60-120 µE/m^2/s in the range 400 – 700nm) all described	Refer to test guidelines	No: test is not adequate
8	pH of test solutions at start and end of test	Refer to test guidelines	No: test is not adequate
9	Method of measuring cell density described	Refer to test guidelines	No: test is not adequate

Item	Requirement	Comment	Concern level
10	Results: Cell density in each flask at each measuring point Mean cell density Growth curve Relationship between concentration of test substance and effects	Refer to test guidelines	No: test is not adequate
11	EC50 values plus method of determination and statistics	Refer to test guidelines	No: test is not adequate
12	Control cell density increased by more than 16 times in 72 hours (i.e. growth rate of 0.9d+1)	Refer to test guidelines	No: test is not reliable
13	The variation coefficient of the control specific growth rates should not exceed 7%	Refer to test guidelines	No: test is not reliable
14	Control pH has not varied by more than 1 unit (i.e. ±0.5 from initial value) during the test	Refer to test guidelines	No: test is not reliable
15	NOEC values plus method of determination and statistics	Refer to test guidelines	No: test is not adequate

A2.7 Part A of the OSPAR Protocols on Methods for the Testing of Chemicals Used in the Offshore Oil Industry- Corophium

Item	Requirement	Comment	Concern level
1	Test organism as listed in Annex B fully description of collection, handling and verification of taxonomy	Refer to test guideline. At a minimum the species, source and cultivation method are required	No test is not adequate
2	Salinity temperature and dO_2 of water near collection site reported	This data may be used to interpret the results from the test with respect to the origin of the organisms	No: Test is not reliable
3	Test sea water quality documented; salinity, pH, dO_2, ammonia nitrogen, nitrite, pesticides and metals	This is a fundamental requirement for any aquatic toxicity test as contamination with metals or pesticides can be a confounding factor in a toxicity test.	No: The toxicity seen may be due to substances other than those added as the test item
4	Acclimatisation and test conditions monitored, reported and within correct limits	These data may be used to interpret the results from the test with respect to the origin of the organisms	No: Test is not reliable
5	Sediment collection, transport preparation and characterisation conducted and reported	These data may be used to interpret the results from the test with respect to the origin of the organisms	No: Test is not reliable
6	Spiking method fully described	Refer to test guideline	No: Test is not adequate
7	Test method and conditions fully described	Refer to test guideline	No: Test is not adequate
8	Raw data; number of mortalities and dates	Refer to test guideline	No: Test is not adequate
9	Calculations, confidence limits and statistical methods fully described	Refer to test guideline	No: Test is not adequate
10	Comment on any other effects or deviations	Refer to test guideline	No: Test is not adequate

A2.8 ISO 14669 Acartia

Item	Requirement	Comment	Concern level
1	Test organism as listed in ISO 14669, full description of collection, handling and verification of taxonomy	Refer to test guideline. At a minimum the species, source and cultivation method are required	No: test is not adequate
2	Temperature 20±2°C, 16h/8h light/dark photoperiod used	Refer to test guideline	No: test is not adequate
3	Test species culture conforms to annex B	Refer to test guideline	No: test is not adequate
4	Test sea water quality documented; Natural water (salinity, pH, dO_2, ammonia nitrogen, nitrite, pesticides and metals) or artificial sea water fully described	This is a fundamental requirement for any aquatic toxicity test as contamination with metals or pesticides can be a confounding factor in a toxicity test.	No: The toxicity seen may be due to substances other than those added as the test item
5	Acclimatisation and test conditions monitored, reported and within correct limits	This data may be used to interpret the results from the test with respect to the origin of the organisms	No: Test is not reliable
6	Preparation and application of test substance fully described	Refer to test guideline	No: Test is not adequate
7	Test method and conditions fully described	Refer to test guideline	No: Test is not adequate
8	Raw data; number of mortalities and dates	Refer to test guideline	No: Test is not adequate
9	Calculations, confidence limits and statistical methods fully described	Refer to test guideline	No: Test is not adequate
10	Was the mortality in controls less than or equal to 10%	Refer to test guideline	No: Test is not reliable
11	Was the toxicity of the reference chemical within the range specified in the ISO document?	Refer to test guideline	No: Test is not reliable
12	Was the dO_2 at the end of the test ≥ 4 mg/L?	Refer to test guideline	No: Test is not reliable
13	Comment on any other effects or deviations	Refer to test guideline	No: Test is not adequate

A2.9 Fish acute toxicity full test modified OECD 203/ Part B of the OSPAR Protocols on Methods for the Testing of Chemicals Used in the Offshore Oil Industry

Item	Requirement	Comment	Concern level
1	Is the mortality in controls less than 10%	Refer to test guideline	No: Test is not reliable
2	Were test conditions semi static and maintained throughout the study	Refer to test guideline	No: Test is not adequate
3	Was dO_2 maintained at ≥60% throughout the test	Refer to test guideline	No: Test is not reliable
4	Were fish held in the lab for at least 12 days with a 12 to 16 hour light photoperiod at 13.5 to 16.5°C (turbot) at ≥60% dO_2 and mortalities acceptable	Refer to test guideline	No: Test is not reliable
5	Was feeding ceased 24hours before treatment	Refer to test guideline	No: Test is not reliable
6	Are exposure conditions fish maximum loading 1g/litre, 12 to 16 hour light photoperiod at 13.5 to 16.5°C (turbot) at ≥60% dO_2, no feeding,	Refer to test guideline	No: Test is not reliable
7	Geometric series of 5 concentration tested with at least 7 Turbot per concentration	Refer to test guideline	No: Test is not adequate
8	For unstable substances the frequency of media changes and of analytical chemistry must be appropriate	Refer to test guideline	No: Test is not adequate
9	Measurement of pH, dO_2, salinity & temperature reported	Refer to test guideline	No: Test is not adequate
10	Daily mortality reported at 24,28, 72 and 96 hours	Refer to test guideline	No: Test is not adequate
11	All procedures fully described in report	Refer to test guideline	No: Test is not adequate
12	Control mortalities reported	Refer to test guideline	No: Test is not adequate
13	Any deviations from protocol reported	Refer to test guideline	No: Test is not adequate

A2.10 Fish acute toxicity limit test modified OECD 203/ Part B of the OSPAR Protocols on Methods for the Testing of Chemicals Used in the Offshore Oil Industry

Item	Requirement	Comment	Concern level
1	Is the mortality in controls less than 10%	Refer to test guideline	No: Test is not reliable
2	Were test conditions semi static and maintained throughout the study	Refer to test guideline	No: Test is not adequate
3	Was dO_2 maintained at ≥60% throughout the test	Refer to test guideline	No: Test is not reliable
4	Were fish held in the lab for at least 12 days with a 12 to 16 hour light photoperiod at 13.5 to 16.5°C (turbot) at ≥60% dO_2 and mortalities acceptable	Refer to test guideline	No: Test is not reliable
5	Was feeding ceased 24hours before treatment	Refer to test guideline	No: Test is not reliable
6	Are exposure conditions fish maximum loading 1g/litre, 12 to 16 hour light photoperiod at 13.5 to 16.5°C (turbot) at ≥60% dO_2, no feeding,	Refer to test guideline	No: Test is not reliable
7	Concentration tested is that which was equivalent to the LC_{50}in the most sensitive of the other species tested with this substance tested with at least 7 Turbot per concentration	Refer to OSPAR guideline	No: Treat result cautiously and consider possibility that the report is not reliable
8	For unstable substances the frequency of media changes and of analytical chemistry must be appropriate	Refer to test guideline	No: Test is not adequate
9	Measurement of pH, dO_2, salinity & temperature reported	Refer to test guideline	No: Test is not adequate
10	Daily mortality reported at 24,28, 72 and 96 hours	Refer to test guideline	No: Test is not adequate
11	All procedures fully described in report	Refer to test guideline	No: Test is not adequate
12	Control mortalities reported	Refer to test guideline	No: Test is not adequate
13	Any deviations from protocol reported	Refer to test guideline	No: Test is not adequate

A2.11 OECD 305 Bioaccumulation

Item	Requirement	Comment	Concern level
1	Is the result a value between 1.5 and 6 (or above) ?	Refer to test guideline	No; test is not adequate
2	If a radiolabelled substance has been used: is the radiochemical purity greater than 95%, is the label in the most stable part of the molecule (with respect to metabolism and degradation), is the specific activity high enough o give suitable sensitivity to the test?	The RCP and specific activity are fundamental to the quality of a radiolabel study.	No: Test is not reliable
3	If a non labelled test substance is used is there a suitable analytical technique for which the procedural recoveries from high and low spiked samples between 70 and 110% ?	A valid analytical method is essential for conducting a study with non-radiolabelled material	No: Test is not reliable
4	Is the water used of constant quality with a pH in the range 6.5 to 8 and not varying by more than ±0.5 pH units?	Refer to test guideline	No: test is not reliable
5	Has the water supply been analysed for heavy metals, cations, pesticides, TOC (<2mg/l) and suspended solids (5mg/l) within the last 3 to 6 months ?	This is a fundamental requirement for any aquatic toxicity test as contamination with metals or pesticides can be a confounding factor in a toxicity test.	No: The toxicity seen may be due to substances other than those
6	Are water characterisation data reported including pH, hardness, total solids, TOC, ammonium, nitrite, alkalinity and for marine species salinity ?	Refer to test guideline	No: test is not adequate
7	Is test substance preparation fully described and appropriate including the use of solvents and dispersion agents ?	Refer to test guideline	No: test is not adequate
8	Has the flow through or semi static system been chosen and set up appropriately ?	Refer to test guideline	No: test is not adequate

Item	Requirement	Comment	Concern level
9	Have fish been acclimatised for 2 weeks at the test temperature and fed with the same food used in the test ?	Refer to test guideline	No: test is not reliable
10	After 48 h settling in period were mortalities at an acceptable level and fish healthy ?	Refer to test guideline	No: test is not reliable
11	Was uptake phase 28 days or was equilibrium demonstrated to have been achieved before that?	If not caution should be exercised when interpreting results. If yes then the kinetic BCF value should be reported	No: Caution required in interpreting data
12	Is the depuration phase half the uptake phase?	This is a minimum value	If less then test is not reliable
13	Were at least 4 fish used per sample time and the loading rate not more than 1g/litre ?	Refer to test guideline	No: test is not adequate
14	Were fish fed at 1 to 2% of body weight with a food of known lipid and protein content ?	Refer to test guideline	No: test is not adequate
15	Were two test concentrations used ?	Refer to test guideline	No: test is not adequate

A2.12 ASTM E1022 Bioaccumulation

No specific guidance is provided for this test protocol here as data based upon this protocol has never been submitted as part of a HOCNF to Cefas at the time of writing. This protocol is however very precise in its wording and the audit requirements for such a study should be based on the definitions cited in paragraph 3.1.1 of this protocol;

"3.1.1 The words "must," "should," "may," "can," and "might" have very specific meanings in this guide. "Must" is used to express an absolute requirement, that is, to state that the test ought to be designed to satisfy the specified condition, unless the purpose of the test requires a different design. "Must" is used only in connection with factors that relate directly to the acceptability of the test (see 13.1). "Should" is used to state that the specified condition is recommended and ought to be met if possible. Although violation of one "should" is rarely a serious matter, violation of several will often render the results questionable. Terms such as "is desirable" are used in connection with less important factors. " May" is used to mean "is (are) allowed to," "can" is used to mean "is (are) able to," and "might" is used to mean "could possibly." Thus the classic distinction between "may" and "can" is preserved, and "might" is never used as a synonym for either "may" or "can."