

Technical University of Denmark



Validation of full life-cycle test with the copepod *Acartia tonsa*

Report to Nordic Council of Ministers and OECD

Kusk, Kresten Ole; Wollenberger, Leah

Publication date:
2005

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

[Link back to DTU Orbit](#)

Citation (APA):

Kusk, K. O., & Wollenberger, L. (2005). Validation of full life-cycle test with the copepod *Acartia tonsa*: Report to Nordic Council of Ministers and OECD. Kgs. Lyngby: Institute of Environment & Resources, Technical University of Denmark.

DTU Library

Technical Information Center of Denmark

General rights

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

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

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Report to
Nordic Council of Ministers
and
OECD

Validation of Full Life-Cycle test with the
copepod *Acartia tonsa*

K. Ole Kusk and Leah Wollenberger

Institute of Environment & Resources
Technical University of Denmark
Kgs. Lyngby, Denmark

2005-09-12

Preface.....	3
Abstract.....	4
1 Introduction.....	5
1.1 Rationale for the Proposed Test Method	5
1.2 Work plan.....	6
1.3 The Prevalidation Testing	7
1.3.1 Chemicals.....	7
1.3.2 Participants in pre-validation work with calanoid copepods	7
2 Method and Materials	8
2.1 Guideline.....	8
2.2 Observed Parameters	9
2.3 Test substances.....	10
2.3.1 3,5-Dichlorophenol (3,5-DCP)	10
2.3.2 Fipronil.....	10
2.4 Statistical treatment of data.....	11
3 Results	11
3.1 Acute Tests with <i>Acartia tonsa</i>	11
3.2 Larval Development Ratio Test on Generation F0	13
3.3 Life-cycle Test with <i>Acartia tonsa</i>	21
3.3.1 Mortality and sex ratio in mass culture	21
3.3.2 Egg Production.....	25
3.3.3 Second generation LDR.....	28
4 Chemical Analysis.....	34
4.1 3,5-Dichlorophenol	34
4.2 Fipronil.....	34
4.2.1 Stock solution.....	34
4.2.2 Water samples	35
5 Discussion and conclusions	37
5.1 Chemical analysis	37
5.2 Test results	37
5.2.1 Acute tests	37
5.2.2 Larval development ratio test on first generation (F0)	37
5.2.3 Life-cycle test.....	37
5.3 Suggested changes and overall evaluation.....	39
5.3.1 General changes	39
5.3.2 Changes of validity criteria.....	40
5.3.3 General conclusion.....	41
References.....	42
6 Annex 1 – Comments.....	44

Preface

This report describes the pre-validation results from *Acartia tonsa* life-cycles tests performed according to the draft OECD guideline with the calanoid copepod *Acartia tonsa* (OECD 2004).

The report was sent out as a draft in May 2005. In the meantime comments have been made by Henrik Tyle, Danish Environmental Protection Agency, and Tim Williams, AstraZeneca UK Limited. These comments as well as suggestions based on a detailed discussion between the authors and Estelle Bjørnstad and Connie Seierø, DHI, are considered in the present report. The comments have as far as possible been responded to directly by answering their mails. The given comments and the responses is included as an Annex (see Annex 1) to this report together with some supporting data from other tests, where the first draft guideline or a method for egg production test close to the one proposed here has been used.

A new version of the draft Guideline on *Acartia tonsa* life-cycle test has been prepared based on the prevalidation work, the comments and the following discussion with DHI.

Abstract

A draft OECD guideline for a life-cycle test with the calanoid copepod *Acartia tonsa* has been prepared. The method was subjected to a pre-validation conducted by two laboratories using 3,5-dichlorophenol and fipronil as reference compounds.

The strong acute toxicity of fipronil prevented to observe any other toxic effect. Moreover, the lowest concentrations of fipronil used in the life-cycle test were below the limit of quantification of the applied analytical method. Hence, it is recommended to use another more suitable reference compound in further validation studies

3,5-DCP, a commonly used reference compound in toxicity testing, was also applied in this study, although it has no endocrine disrupting potential. Unfortunately, fipronil was also not suitable for detection of endocrine disrupting effects due to its very high acute toxicity

The part of the draft guideline considering egg production needs to be revised to secure a proper number of egg producing females and thus to improve the quality of the results from this part of the guideline.

The validation study revealed that the LDR at the end of the LDR test should be in the range 50 ± 20 % to obtain results with adequate precision. Hence, we propose to include this as a new validity criterion for the test.

The base for evaluation of precision is limited. Only for the LDR (larval development ratio) test on the first and second generation with 3,5-DCP there were two data sets available allowing comparison of results. The EC50 determined by the two laboratories differed only by a factor of two to three and can thus be considered to be in the same range.

An increase in the number of replicates in the LDR tests from 4 to 6 will be suggested as an option in the revised draft guideline to improve the precision of the results and reduce the confidence limits.

Other validity criteria need to be revised:

- The suggested survival ratio in the control mass culture of at least 80 % has been proven to be too strict since *Acartia* is a short-lived organism with a high reproductive output and with a natural mortality well above 20 % within three weeks. In the present study mortality ranged between 50 and 60 %. More experimental work is required to define this validity criterion properly.
- The suggested control sex ratio of 40-60 % can often not be met. Many factors influence it and hence, it is suggested to leave out this criterion for the time being. More research is needed to identify factors influencing the sex ratio in *Acartia tonsa*.
- To replace the two validity criteria mentioned above a new validity criterion is suggested: at least 50 living males and 50 living females must be present in the control at the end of the life-cycle test.
- The validation study has confirmed that development of *A. tonsa* responses very sensitive to changes in temperature. Therefore, the test temperature shall be 20 ± 1 °C

Based on the results of the pre-validation study the existing draft guideline will be revised. The revised guideline will be ready for further validation in an interlaboratory exercise.

1 Introduction

1.1 Rationale for the Proposed Test Method

Copepods are common organisms in freshwater as well as marine areas. They constitute the food and energy link between primary producers and higher-level heterotrophic organisms such as bigger crustaceans and fish larvae. Thus, they play an important role in aquatic ecosystems.

Pelagic and benthic copepods are abundant and common in most aquatic environments. The calanoid copepods dominate the marine and brackish zooplankton in coastal areas and to a lesser extent also the open sea, where they graze the phytoplankton. The harpacticoids are abundant on and in the upper-most layers of the sediment, where they graze on bacteria growing on organic materials. Both groups of copepods, thus, convert micro-organisms into multicellular tissue and thereby contribute to the energy transfer from primary producers to higher trophic levels in the ecosystems

Copepods have a sexual reproduction. A female is fertilised by a male, which place a spermatophore close to the female genital opening. The spermatozoa may be stored in the female's reproductive system and fertilise eggs produced over a longer time period (Mauchline 1998).

Most calanoid copepods release their eggs singly in the water as for example *Acartia*, but some produces egg sacs where eggs are stored until hatching.

Copepods have been studied for many years and also been cultured in the laboratory for toxicity testing and as food for fish larvae.

A standard method (ISO 1997) for acute toxicity testing with marine copepods exists since 1997. Three species are used as test organisms – these are *Acartia tonsa*, *Tisbe battagliai* and *Nitocra spinipes*.

At the moment there are no internationally harmonized (i.e. OECD) chronic test methods for marine invertebrates, although they are an ecologically important and large group of organisms that need to be protected. Therefore, the development of a test on reproduction and development of marine copepods has been taken on the work plan of the OECD Test Guideline Program. The copepod life-cycle test is urgently needed in several regulatory systems. It will be of wide applicability and as such facilitate international harmonization of risk and hazard assessments. It addresses important endpoints and environmental compartments (marine/estuarine/brackish) that are poorly covered by existing methods. Four common and regularly used species in toxicity tests are proposed as model test organisms, namely the calanoid copepod *Acartia tonsa* and the harpacticoid copepods *Nitocra spinipes*, *Amphiascus tenuiremis* and *Tisbe battagliai*. The new guidelines are intended for evaluation of adverse long-term effects of various types of chemicals to aquatic invertebrates – including

compounds used in off shore oil industry, pharmaceuticals and endocrine disrupters as well as industrial chemicals.

The guideline evaluated in the present report, therefore, has a great potential and contributes to filling gaps in the existing test battery.

1.2 Work plan

In June 2003 a Nordic Expert group under NordUtte prepared a proposal for a new guideline “OECD draft Guideline for Testing of Chemicals – Copepod Development and Reproduction Test”. This guideline describes testing of effects on development and reproduction of marine calanoid and harpacticoid copepods in full life-cycle tests. The method is intended also to cover effects of potential endocrine disrupters.

This draft guideline was discussed at the first OECD meeting of the ad hoc expert group on invertebrate testing in November 2003 in Paris (OECD 2003). Differences in biology leading to differences in methods, especially with regard to handling of animals, feeding and media, made it difficult to include both harpacticoid and calanoid copepods in one common guideline. It was therefore decided to split the first draft into two – one with the benthic species (*Nitocra spinipes*, *Amphiascus tenuiremis* and *Tisbe battagliai*) and one with the pelagic species *Acartia tonsa*. It was further decided that the drafts should be ready by the end of 2003 and be commented within the first two months of 2004. A pre-validation test should start in spring 2004. OECD planned to finalise the pre-validation work by June 2004 followed by an interlaboratory study on the two draft guidelines. Unfortunately, the suggested time schedule could not be met.

The work plan for the present project as outlined in the contract between Environment & Resources DTU and NordUtte was as follows:

- *Prepare a new draft guideline on *Acartia tonsa* (Nov-Dec 2003)*
- *Collect and compile comments on the revised draft and include where relevant (Jan 2004)*
- *Plan and co-ordinate (in co-operation with the Swedish partner) the validation ring-test (Feb-March 2004)*
- *Perform validation tests with one or two of the proposed reference compounds according to the revised draft (Feb-April 2004)*
- *Report the outcome of the validation to NordUtte and OECD (April-May 2004)*

Preparation of proposals for TGs with copepods

Revised drafts for the two test guidelines were prepared in December 2003 by Magnus Breitholtz, Stockholm University, and K. Ole Kusk and Leah Wollenberger, Technical University of Denmark, respectively, and sent out for commenting in January 2004 by OECD.

The received comments on the method with calanoids were compiled and responded to in Marts 2004.

The comments gave rise to a few changes in the calanoid draft guideline – the main one being the exclusion of natural water as media, since synthetic media has been used successfully for years for cultivation of and toxicity testing with *A. tonsa* (Kusk and Wollenberger 1999).

Other comments considered suggested a harmonisation in the structure of both TGs. Based on the comments a (final) Draft Document “OECD Draft Guidelines for Testing of Chemicals. Proposal for a new guideline. Calanoid Copepod Development and Reproduction Test with *Acartia tonsa*” was prepared and sent to OECD in March 2004 (OECD 2004).

1.3 The Prevalidation Testing

1.3.1 Chemicals

At the expert meeting in OECD November 2003 (OECD 2003) it was decided to use two chemicals in the prevalidation work and perform the full test with both of them. The chemicals were:

- 3,5-dichlorophenol (DCP)
- fipronil

3,5-dichlorophenol is often used as a reference compound in various standardised toxicity test methods. It is water soluble, not ready biodegradable, non-volatile, and has a relatively low K_{ow} . This means that the compound is easy to handle and is expected to stay in the water phase for the duration of the test period.

Fipronil is an insecticide suspected to be an endocrine disrupter (Cary et al. 2004). It is acutely toxic at the low ppb level ($LC_{50} < 10 \mu\text{g/L}$ for crustaceans) (Connelly 2001). It has a water solubility of 20 mg/l, a K_{ow} of 4.01 and an aerobic aquatic half-life of 14.5 days (Connelly 2001), which means that there may occur adsorption on glass and food as well as biodegradation.

Physical-chemical properties of the two chemicals are included in chapter 2.3.1.

1.3.2 Participants in pre-validation work with calanoid copepods

Three laboratories agreed to participate in the prevalidation work with *Acartia tonsa*. These are:

- Environment & Resources DTU, Kgs. Lyngby, Denmark (K. Ole Kusk)
- DHI Water and Environment, Hørsholm, Denmark, (Estelle Bjørnestad)
- ECT Oekotoxikologie GmbH, Floersheim, Germany, (Michael Meller)

ECT Oekotoxikologie GmbH (ECT) decided relatively late to join the prevalidation work. This laboratory had no experience with *Acartia*, and Michael Meller visited E&R as well as DHI Water and Environment (DHI) in August 2004 for a short training course and for collection of organisms.

Data sets on tests with *Acartia tonsa* and the test chemicals fipronil and 3,5-dichlorophenol have been received from DHI. From ECT only data on an acute test with *Acartia tonsa* and 3,5-dichlorophenol has been received.

At DTU the prevalidation work started in March 2004 and continued until July 2004. Tests with both reference chemicals were conducted according to the Draft Guideline for the Calanoid copepod *Acartia tonsa* (March 2004 version).

Chemical analyses for validation of nominal fipronil concentrations were performed by SOFIA GmbH, Berlin, Germany, for water samples from DTU.

2 Method and Materials

2.1 Guideline

The test was performed as described in (Final) Draft Document “OECD Draft Guidelines for Testing of Chemicals. Proposal for a new guideline. Calanoid Copepod Development and Reproduction Test with *Acartia tonsa*. March 2004.”

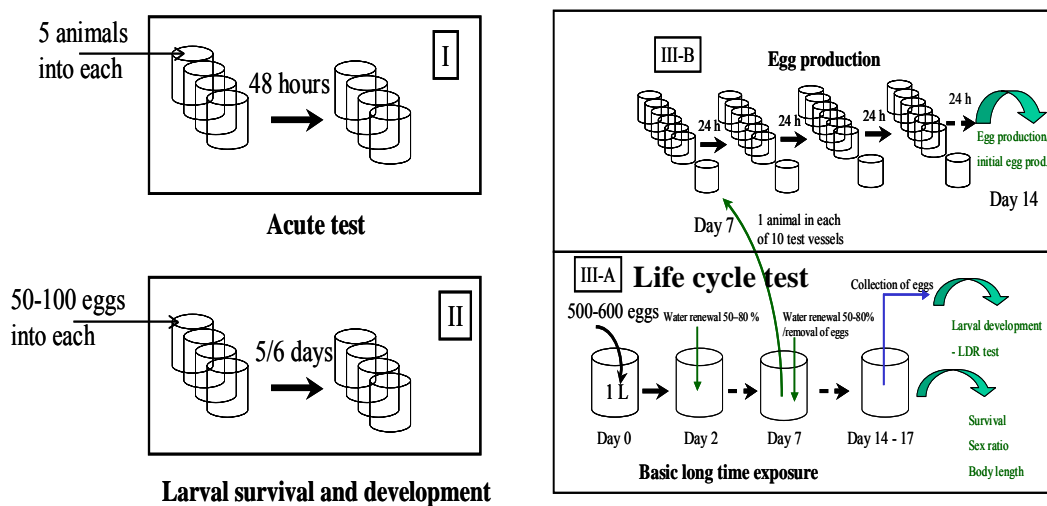


Figure 1: Overview of the different parts of the test. This figure is later used to illustrate to which part of the procedure the text refers.

DHI made some modifications to the test guideline. These were:

- Use of natural saltwater with a salinity of 32 ‰
- The LDR tests were stopped after exactly 120 hours.
- The egg production test was started with groups of 14 days old females

These modifications have probably influenced the test results.

2.2 Observed Parameters

Figure 1 gives an overview of the different components of the life-cycle test that have been performed in this pre-validation work. To help keeping an overview this figure is later used to illustrate to which part of the procedure the text refers by indicating the actual procedure with a red frame.

Acute toxicity (I) - actually not part of the guideline but gives essential information for selection of test concentrations

LDR (II) - Larval Development Ratio (Early Life stages development) on the **F0 generation** – also this part was not included in the guideline, but again, it gains information important for the selection of test concentrations in the life-cycle exposure.

Life cycle test (III A and B)

Control of hatching to secure that the quality of eggs is satisfactory (validity criteria is hatching $\geq 80\%$).

Maturity - Observation of age at first egg production of individually isolated females

Egg production - Number of eggs per female and day of individually isolated animals (validity criteria is ≥ 30 eggs/female per day at day 12 and later in the control)

Size - length of animals

Mortality of individually **isolated control animals** in the egg production part (validity criteria $\leq 20\%$ mortality)

Mortality in basic exposure control (unexposed) (validity criteria $\leq 20\%$ mortality) measured at the end of the exposure (14-17 days)

LDR - Larval Development Ratio (Early Life stages development) on the **F1 generation**

Control mortality in LDR test (validity criteria $\leq 20\%$ control mortality)

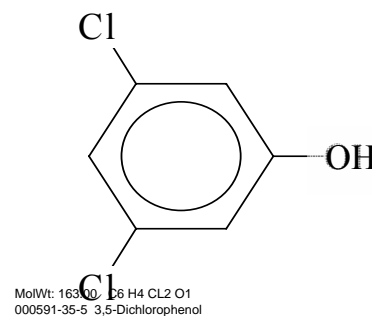
Sex ratio (validity criteria is between 40 and 60 % of each sex in the control)

Other parameters: Egg size of fipronil-exposed females

2.3 Test substances

2.3.1 3,5-Dichlorophenol (3,5-DCP)

DTU and DHI used the same batch of 3,5-DCP: Riedel-de Haën Product No 47070 (PESTANAL) UN 2020 Lot 2262X (99.8 %)



Physical-chemical properties:

CAS Number:	591-35-5
Molecular weight:	163.0
Water solubility:	5380 mg/L (25 °C) (Huyskens, P et al.1975)
MP (deg C):	68 (Hansch, C et al. 1995)
BP (deg C):	233 (Hansch, C et al.. 1995)
Log Kow :	3.62 (Hansch, C et al. 1995)
Vapor Pressure :	0.00842 mm Hg (extrapol.) (25 °C) (Shiu, WY et al.1994)
pKa :	8.36 (SPARC on-line calculator)

Toxicity data: (US EPA ECOTOX database except otherwise indicated)

<i>Acartia tonsa</i>	48 h LC50: 0.51 mg/L
<i>Acartia tonsa</i>	48 h LC50: 0.95 mg/L (ISO 14669)
<i>Acartia tonsa</i>	48 h LC50: 1.00 mg/L (Bjørnstad et al. 1993)
<i>Crangon septemspinosa</i>	96 h LC50: 1.50 mg/l
<i>Carassius auratus</i> (Goldfish)	96 h LC50: 3-5 mg/L
Zebra fish	24 h LC50: 2.9 mg/L (Zarorc-Končan et al 2002)
<i>Tisbe battagliai</i>	24 h LC50: 10.74 µM (1.75 mg/L)
<i>Daphnia magna</i>	48 h EC50: 2.48 mg/L (Zarorc-Končan et al 2002)

Fate data:

Test for ready biodegradability: 5% degradation in 28 days. (Zarorc-Končan et al 2002)

2.3.2 Fipronil

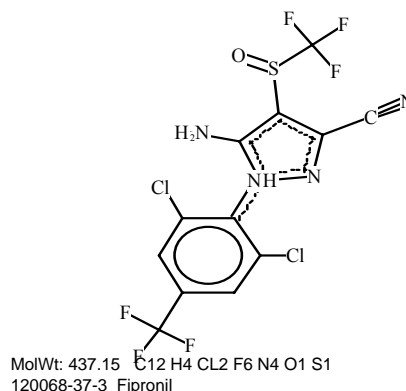
DTU and DHI used the same batch of fipronil: Riedel-de Haën Product No 46451 (PESTANAL) UN 02588 Lot 2218X (97.5 %)

Physical-chemical properties (Connelly 2001):

CAS Number:	120068-37-3
Molecular weight:	437.2
Water solubility:	2.2-2.4 mg/L
Vapor Pressure:	3.7×10^{-4} mPa (25 °C)
Henry's Constant:	3.7×10^{-5} Pa m ³ /mol
Log K _{ow}	4.01

Fate data (Connelly 2001):

Aerobic aquatic half-life:	14.5 days
Field Dissipation half-life:	630-693 days



Toxicity data (Connelly 2001):

Fish	96 h LC50:	0.085-0.248 mg/L (three species)
<i>Daphnia magna</i>	48 h LC 50:	0.19 mg/L
Mysid Shrimp	96 h LC50:	0.00014 mg/L

2.4 Statistical treatment of data

EC₁₀ and EC₅₀ (concentrations reducing the specific observation parameter by 10 and 50%, respectively and their confidence limits were estimated by fitting the obtained continuous response data to the logarithmic normal distribution function. The obtained observation parameter was normalised by dividing by the control response estimate, which here, by contrast to tests with quantal responses, has a non-zero value and therefore a variability that affects the normalized response. The influence of the covariance within the control was taken into account in the statistical calculations using a Taylor expansion of the data followed by inverse estimation (Andersen, 1994). Further, experimental designs were made with more than six control replicates and narrow spacing (factor of maximally three) of test concentrations at all effect levels to give enough data points to allow the estimation of an accurate concentration-response relationship.

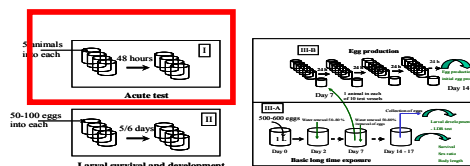
Statistical differences between groups were tested by use of a student T-test.

3 Results

3.1 Acute Tests with *Acartia tonsa*

Results of the acute tests performed according to the ISO 14669 are presented in Tables 1-4. Acute tests are performed to find the relevant concentration range

for the following LDR test on first generation animals and in addition for the sake of the reference compound 3,5-DCP to check the sensitivity of animals.



DTU results:

Table 1: 24 hours Lethal Concentration (LC) for 10 and 50 % of the *Acartia tonsa* population exposed to 3,5-DCP and fipronil. 95% confidence limits in brackets.

Compound	24 h LC10	24 h LC 50
3,5-DCP (mg/L)	0.69 (0.46-0.84)	1.1 (0.92-1.3)
Fipronil (µg/L)	0.60 (0.19-1.1)	11 (6.2-32)

Table 2: 48 hours Lethal Concentration (LC) for 10 and 50 % of the *Acartia tonsa* population exposed for 3,5-DCP and fipronil. 95% confidence limits in brackets.

	48 h LC10	48 h LC 50
3,5-DCP (mg/L)	0.57 (0.22-0.74)	0.95 (0.78-1.1)
Fipronil (µg/L)	0.25 (0.078-0.49)	2.4 (1.5-3.7)

DHI results:

Table 3: 48 hours Lethal Concentration (LC) for 50 % of the *Acartia tonsa* population exposed to 3,5-DCP and fipronil. 95% confidence limits in brackets.

Compound	Test identification	48 h LC 50
3,5-DCP (mg/L)	2003 – March	1.6 (1.3-2.4)
3,5-DCP (mg/L)	2003 – July	1.4 (1.2-1.7)
3,5-DCP (mg/L)	2003 – November	1.3 (1.1-1.5)
3,5-DCP (mg/L)	2004 – February	1.7 (1.5-2.1)
3,5-DCP (mg/L)	2004 – July	1.9 (1.6-2.3)
3,5-DCP (mg/L)	2004 – October	1.1 (0.96-1.4)
3,5-DCP (mg/L)	2005 – January	1.4 (1.2-1.7)
	48 h LC10	48 h LC 50
Fipronil (µg/L)	0.23 (0.09-0.40)	2.5 (1.5-3.6)

ECT results:

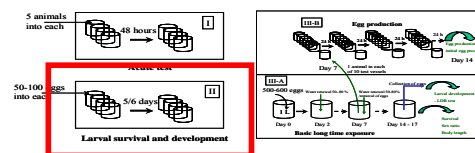
Table 4: 24 and 48 hours Lethal Concentration (LC) for 50 % of the *Acartia tonsa* population exposed to 3,5-DCP. 95% confidence limits in brackets.

Compound	24 h LC50	48 h LC 50
3,5-DCP (mg/L)	1.8 (1.6-2.0)	0.75 (0.64-0.88)

The found 48 h LC50 values for 3,5-DCP are close to the one given in the ISO standard (0.95 mg/L with a repeatability of ± 0.22 mg/L and a reproducibility of ± 0.55 mg/L).

It can be concluded that the sensitivity of *Acartia tonsa* in acute tests is in the same range at the involved laboratories.

3.2 Larval Development Ratio Test on Generation F0



3,5-Dichlorophenol: DTU:

Figure 2 shows the results of the LDR test with 3,5-DCP conducted at DTU. No effects were observed on hatching success, which generally was well above 80 % (average 92.5 %), or on larval survival (generally about 70 %). The control LDR was close to 50 %, which is the optimal control value. At 100 $\mu\text{g/L}$ and higher concentrations a significantly reduced LDR was observed. The validity criterion for hatching (80 %) is fulfilled, whereas the validity criterion for larval control survival (80 %) is not fulfilled.

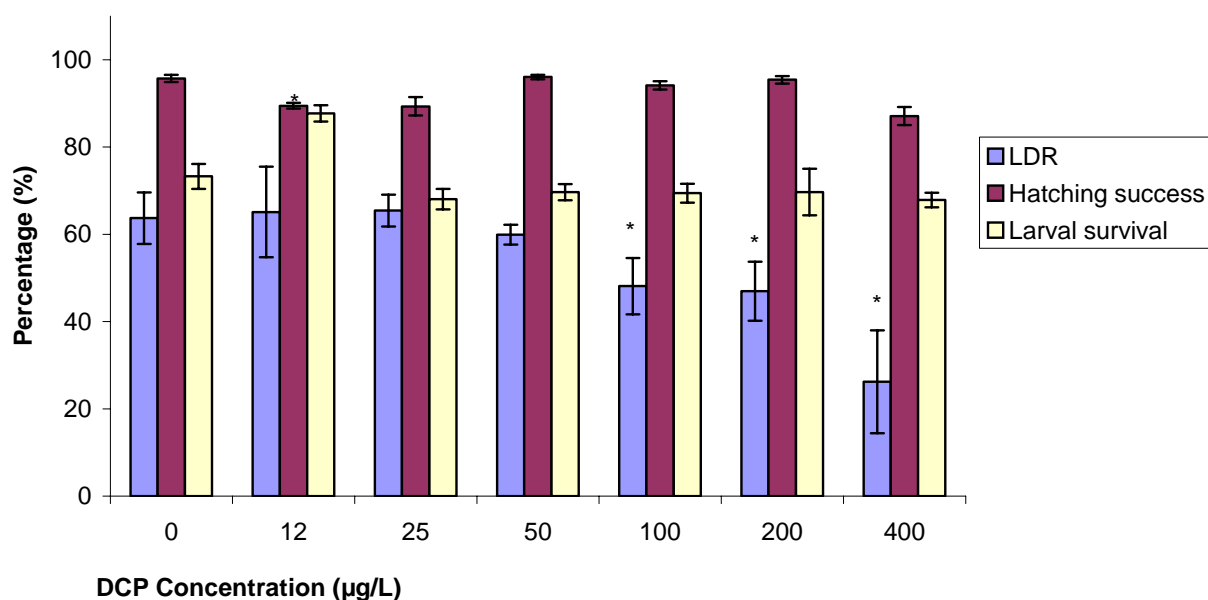


Figure 2: Larval development ratio (LDR = percentage of animals that reached a copepodite stages), hatching success and larval survival of *Acartia tonsa* exposed for 116 hours to 3,5-dichlorophenol in the concentration range from 12 – 400 $\mu\text{g/L}$ at DTU. * indicates significant difference versus control at a 95 % confidence level.

Figure 3 shows the concentration-response curve for the Inhibition of LDR. The estimated EC_{50} was 351 $\mu\text{g/L}$ and the corresponding EC_{10} was 96 $\mu\text{g/L}$ (see Table 5).

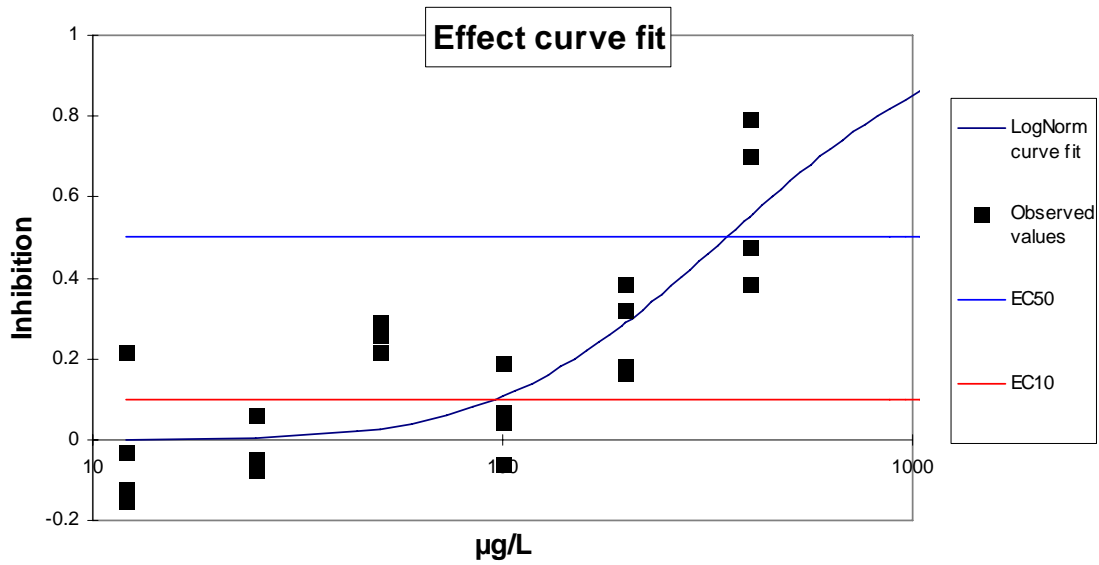


Figure 3: Concentration response curve for LDR test with 3,5-DCP and *Acartia tonsa* exposed for 116 hours at DTU.

Table 5. Calculated EC values with 95% confidence limits for the LDR test with 3,5-DCP and *Acartia tonsa* at DTU. Concentrations are nominal.

		95 % confidence limits	
EC	µg/L	Lower (µg/l)	Upper (µg/l)
10	96	69	134
50	351	241	511

In an earlier study at DTU an EC10 and EC50 for 3,5 DCP of 82 µg/L and 179 µg/L, respectively, were found (Andersen et al. 2001).

3,5-Dichlorophenol: DHI

Figure 4 illustrates effects on LDR, hatching success and larval survival at DCP concentrations in the range from 66 $\mu\text{g/L}$ to 750 $\mu\text{g/L}$. All three parameters were significantly reduced at the three highest concentrations. LDR was also significantly reduced at 220 $\mu\text{g/L}$. One problem though is that the control LDR was very low (< 10 %), which means that more than 90 % of the test animals still were in the nauplii stages. This causes a large uncertainty of the results as is also seen in the concentration-response curve for the LDR in Figure 5. An inhibition of 100 % means that no copepodites were found in the respective replicate. At all concentrations one or more replicates contained no copepodites. To overcome this problem, DHI divided the nauplia larvae into two groups consisting of NI nauplii (nauplii stages I and II) and NII nauplii (nauplii stages III-VI) (see Table 4). Considering the percentage of NII nauplii in relation to the number of eggs applied, the concentration-response curve shown in Figure 6 was achieved. The EC50 values of both calculation methods (percentage of copepodites as well as NII nauplii) are actually quite close: 132 $\mu\text{g/L}$ and 192 $\mu\text{g/L}$, respectively (Table 6 and 8) and both in the same range as the EC50 obtained in the parallel experiment at DTU, which was 350 $\mu\text{g/L}$ (Table 5). The value of 132 $\mu\text{g/L}$ though is very uncertain and is not considered as valid.

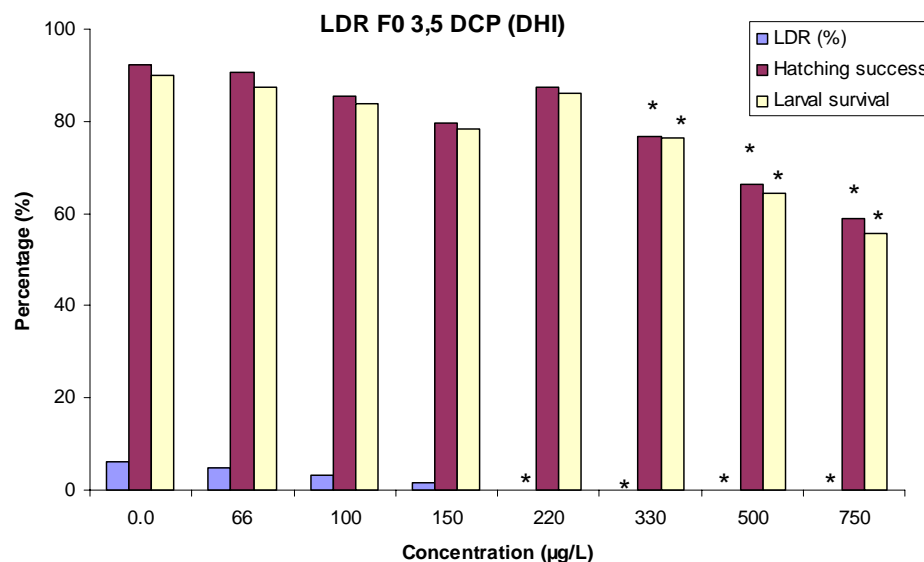


Figure 4: Larval development ratio (LDR = percentage of animals that reached a copepodite stage), hatching success and larval survival of *Acartia tonsa* exposed to 3,5-dichlorophenol in the concentration range from 66 – 750 $\mu\text{g/L}$ at DHI. * indicate significant difference versus control at a 95 % confidence level.

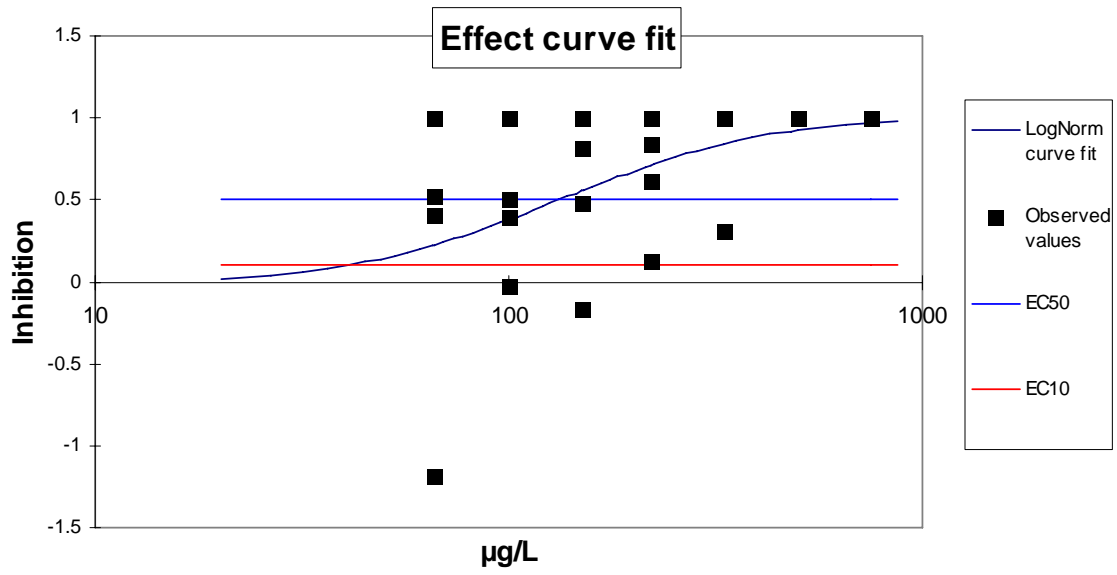


Figure 5: Concentration response curve for LDR test with 3,5-DCP and *Acartia tonsa* at DHI.

Table 6. Calculated EC values with confidence limits for the LDR test with 3,5-DCP and *Acartia tonsa* at DHI. Concentrations are in µg/L (nominal concentrations)

		95 % confidence limits	
EC	µg/L	Lower (µg/L)	Upper (µg/L)
10	41	9	180
50	132	66	266

Table 7: Inhibition of *Acartia tonsa* NII-larval development with 3,5-DCP (DHI)
 Fraction of NII-larvae (in % of total number of added eggs)

Control values

Concentration (mg/l)	NII larvae (% of eggs added)	Inhibition (%)
Control 1	66.0	-
Control 2	72.2	-
Control 3	62.5	-
Control 4	60.0	-
Control 5	57.6	-
Control 6	61.0	-
Control mean	63.2	0

Exposed groups

Concentration (mg/l)	NII larvae (% of added eggs)	Inhibition (%)
0.066	58.3	8
0.066	49.1	22
0.066	61.7	2
0.066	52.7	17
0.100	52.0	18
0.100	49.0	22
0.100	29.5	53
0.100	54.8	13
0.150	32.0	49
0.150	41.0	35
0.150	29.1	54
0.150	62.5	1
0.220	13.7	78
0.220	30.0	53
0.220	13.6	78
0.220	30.6	52
0.330	15.1	76
0.330	18.8	70
0.330	10.8	83
0.330	21.6	66
0.500	9.5	85
0.500	11.8	81
0.500	11.1	82
0.500	7.1	89
0.750	0.0	100
0.750	2.0	97
0.750	0.0	100
0.750	1.8	97

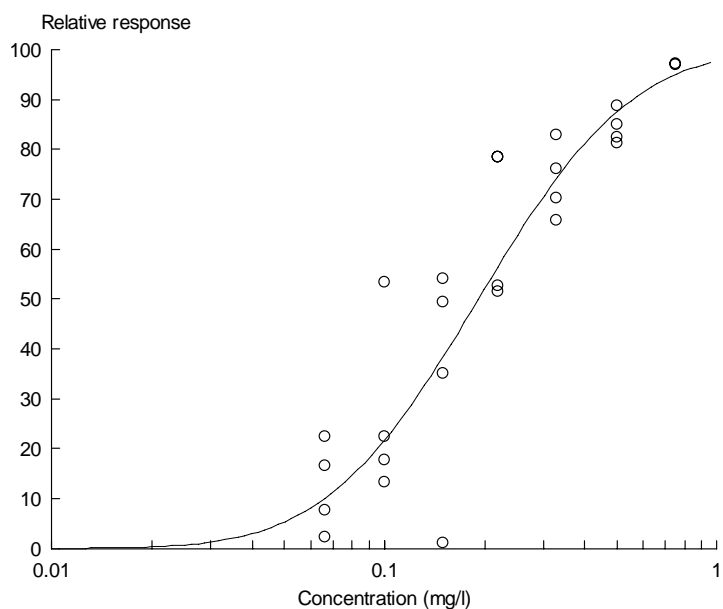


Figure 6: Concentration response curve for inhibition of fraction of NII nauplii of total number of added eggs in test with 3,5-DCP and *Acartia tonsa* at DHI.

Table 8. Calculated EC values and confidence limits for the LDR test with 3,5-DCP and *Acartia tonsa* performed at DHI. Concentrations are nominal. Results are calculated by use of the DHI statistical program Toxedo.

		95 % confidence limits	
EC	µg/L	Lower (µg/L)	Upper (µg/L)
10	66	35	92
50	192	155	236

Fipronil - DTU

The result of the F0 generation LDR test with fipronil is shown in Figure 7. At 37.5 ng/L (nominal concentration) a significant decrease of the LDR occurred, which was not observed at higher and lower concentrations. The decrease at 37.5 ng/L thus seems to be caused by experimental variations rather than by the test compound. Larval survival was as in the experiment with 3,5-DCP at DTU about 70 %, and control hatching success was also high (about 90 %). At two of the three highest concentrations, which were in the acutely toxic range, hatching was weak but significantly reduced.

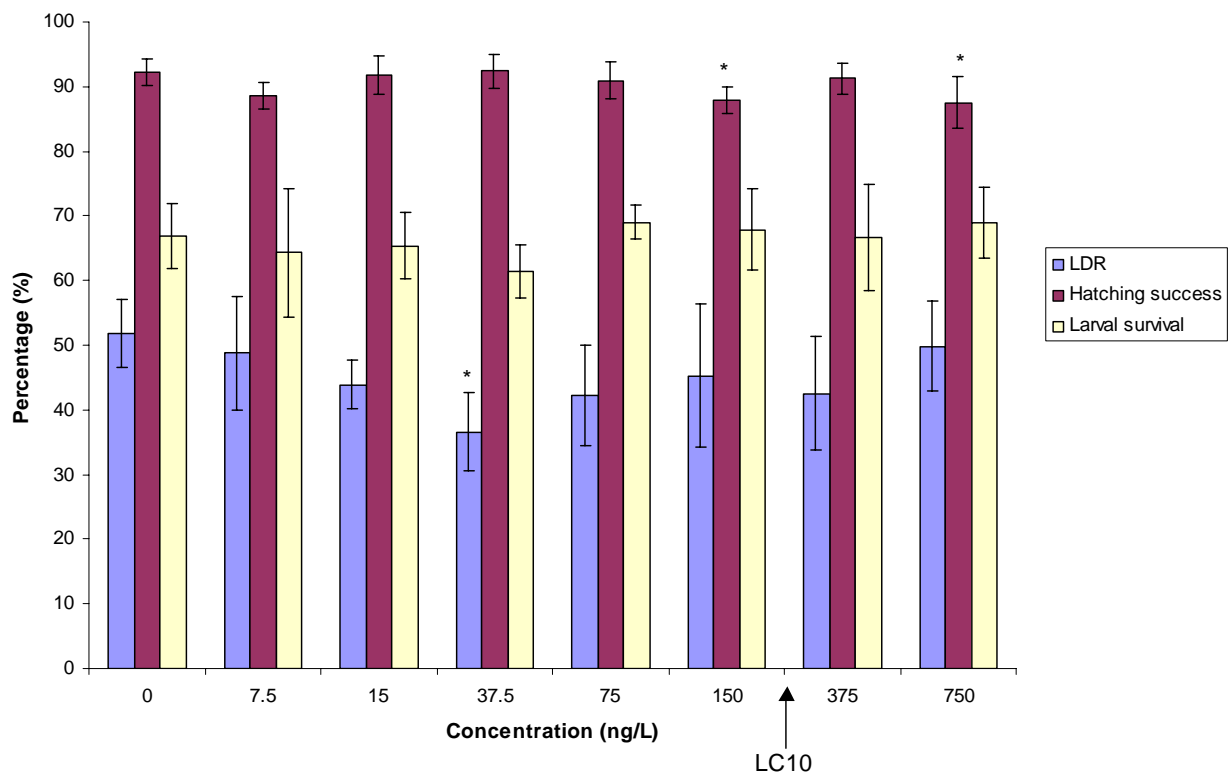


Figure 7: Larval development ratio (LDR = percentage of animals that reached a copepodite stage), hatching success and larval survival of *Acartia tonsa* exposed for 125 hours to fipronil in the concentration range from 7.5 – 750 ng/L at DTU.

* Asterisk indicates significant difference from control (p-value < 0.05).

Fipronil - DHI

In Figure 8 is shown a range finding test result with Fipronil conducted at DHI. The factor between concentrations was 10 in order to cover a broad concentration range. The highest concentrations were thus in the range that also caused effects in the 48 h acute test. A significant reduced survival was seen at two of the three highest concentrations – though not as strong as might be expected from the LC values (Tables 2 and 3), but it should be noted that in contrast to the acute test the animals here were fed with microalgae, which may change their sensitivity and the bioavailability of the test compound.

From Figure 8 is also seen that there was no effect on hatching of *Acartia* eggs. At the highest concentration no copepodites developed and at lower concentrations no significant effects were seen on LDR.

In the concentration range covered by both DTU and DHI (up to 750 ng/L) there were only minor effects - in the DTU experiment on hatching success and in the DHI experiment on survival.

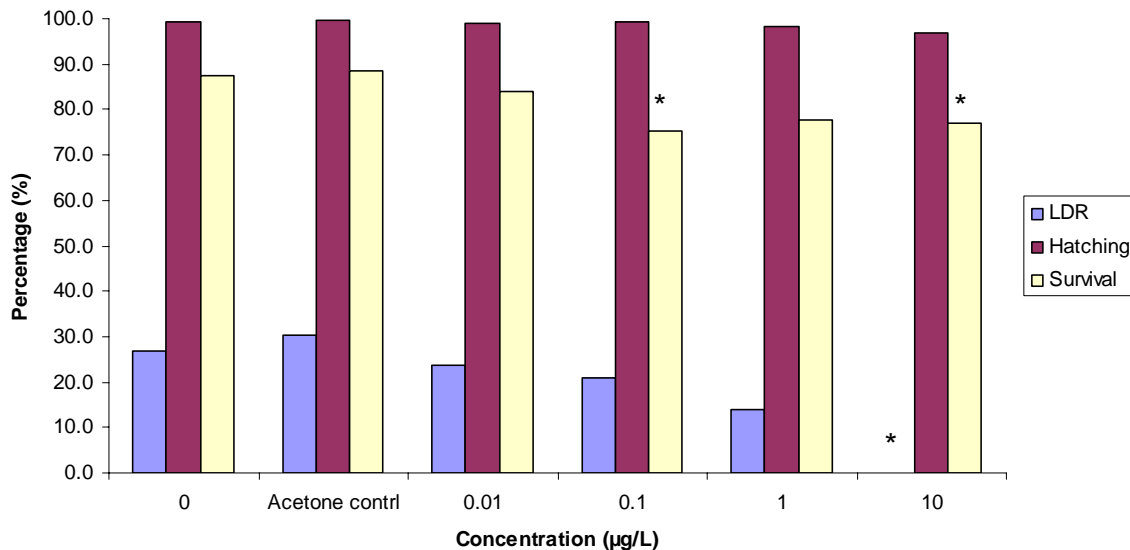
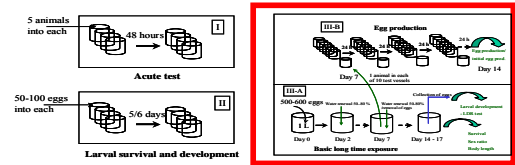


Figure 8: Larval development ratio (LDR), hatching success and larval survival of *Acartia tonsa* exposed to fipronil in the concentration range from 10ng/L – 10 µg/L (DHI). * Asterisks indicate significant difference from control (p-value < 0.05).

3.3 Life-cycle Test with *Acartia tonsa*



Hatching control:

At DTU, four replicates with 40-45 eggs were put up for hatching control. After 3 days the hatching success was measured. The average hatching success was 82 %. The validity criterion for this parameter is ≥ 80 %.

3.3.1 Mortality and sex ratio in mass culture

3,5-DCP and fipronil:

Tables 9 and 10 show the numbers of eggs at the start of the test, and the total number of animals as well as the sex ratio at the end of the test conducted at DTU.

The mortality was between 50 and 60 % in the controls and at the lowest concentrations in tests with both 3,5-DCP and fipronil. The control sex ratios in the two experiments were deviating more than the draft method prescribes (At present the percentage of each sex should be 40-60 %). This is also illustrated in Figure 9 and 10.

At DHI was observed a female ratio of just 40 and 45 % in control and acetone control, respectively. The mortality was just above 50 % in the control whereas it was only 25 % in the acetone control (Figure 11). Thus, the validity criterion for survival in control mass cultures was not fulfilled in any of the experiments at DTU and DHI.

Table 9: Number of eggs at the start of the long-term exposure to 3,5-DCP and number of surviving *Acartia tonsa* after 14 days at DTU. Yellow background indicates deviation from the validity criterion.

Treatment	Initial number of eggs	Ratios		Total number of animals	Mortality %
		Females	Males		
Control	540	0.37	0.63	258	52.2
2.5 µg/L	532	0.54	0.46	226	57.5
7.4 µg/L	533	0.43	0.57	269	49.5
22 µg/L	563	0.52	0.48	173	69.3
67 µg/L	608	0.41	0.59	312	48.7
200 µg/L	535	0.39	0.61	229	57.2

Table 10: Number of eggs at the start of the long-term exposure to fipronil and number of surviving *Acartia tonsa* after 15 days at DTU. Yellow background indicates deviation from the validity criterion.

Treatment	Initial number of eggs	Ratios		Total number of animals	Mortality %
		Females	Males		
Control	547	0.46	0.54	236	56.9
C+acet	484	0.43	0.57	282	41.7
16 ng/L	519	0.37	0.63	234	54.9
40 ng/L	556	0.40	0.60	285	48.7
100 ng/L	490	0.42	0.58	319	34.9
250 ng/L	652	0.34	0.66	385	41.0
750 ng/L	540	0.41	0.59	296	45.2

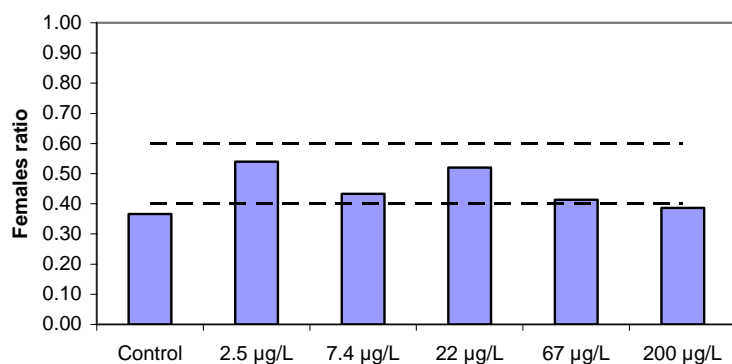


Figure 9 Sex ratio (females) in mass cultures exposed for 14 days to 3,5-DCP at DTU. Dashed lines indicate the range of the validity criterion (40-60 % females).

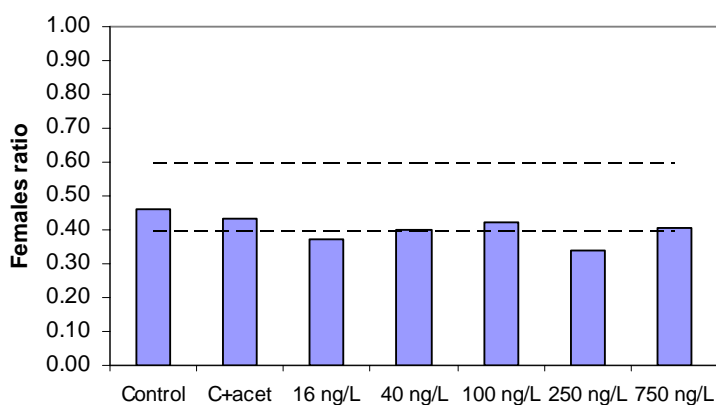


Figure 10 Sex ratio (females) in mass cultures exposed for 15 days to fipronil at DTU. Dashed lines indicate the range of the validity criterion (40-60 % females).

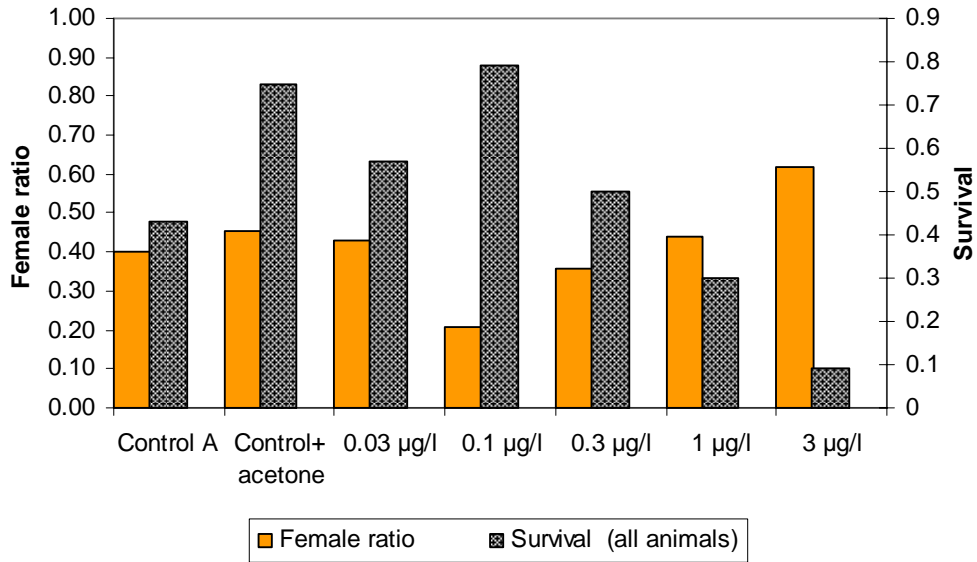


Figure 11. Sex ratio (females) and survival in mass cultures exposed to fipronil at DTU.

Body length:

At DTU a significantly increased body length of animals exposed to 200 mg/L 3,5-DCP was observed. No effect on this parameter was observed at lower concentrations (Figure 12).

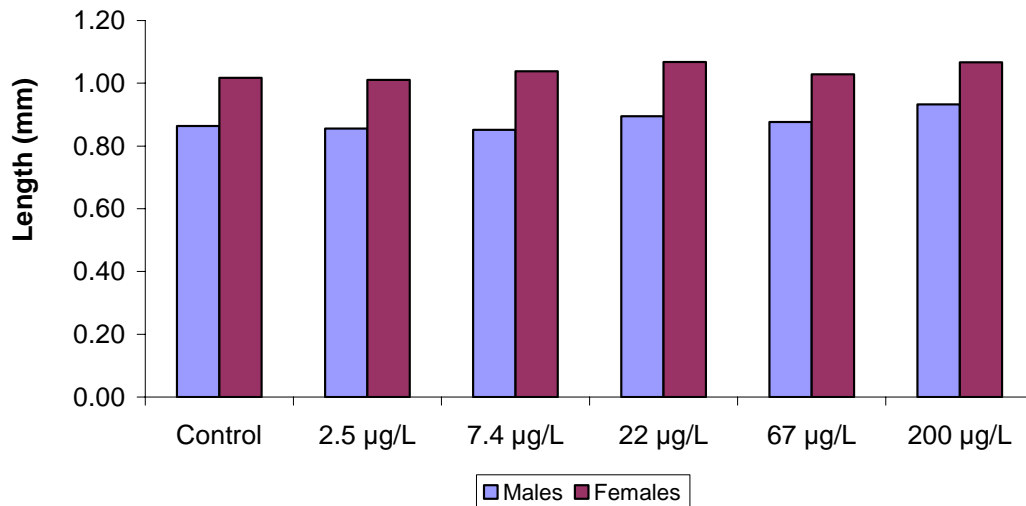


Figure 12. Average body lengths of males (n=25) and females (n=25) in mass cultures of *Acartia tonsa* exposed for 14 days to 3,5-DCP at DTU

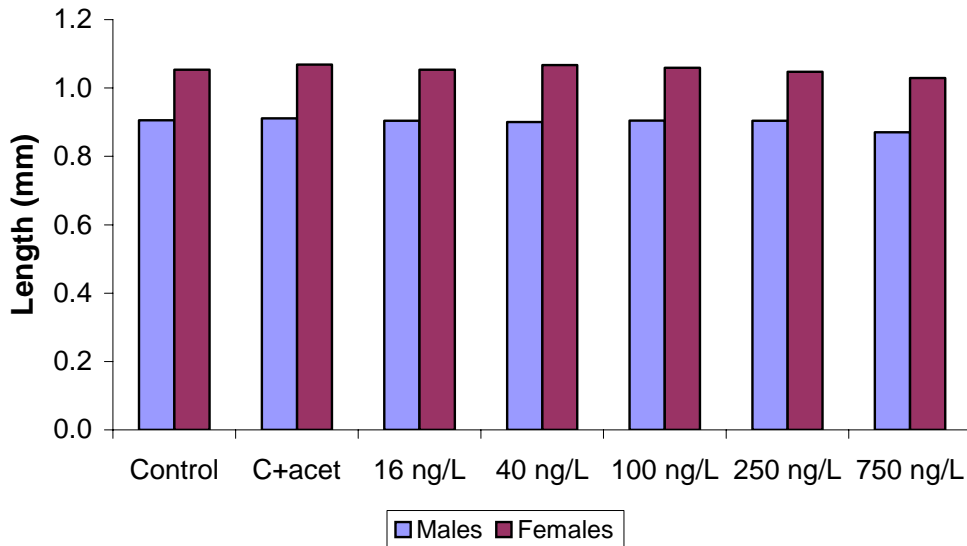


Figure 13 Average body lengths of males (n=25) and females (n=25) in mass cultures of *Acartia tonsa* exposed for 15 days to **fipronil** at DTU.

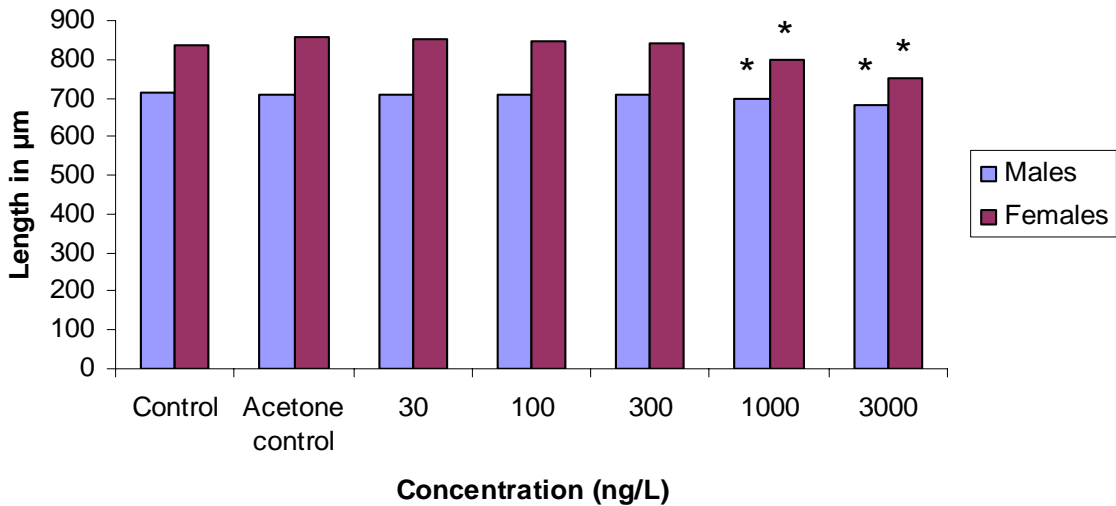


Figure 14 Average body lengths of males (n=25) and females (n=25) in mass cultures of *Acartia tonsa* exposed for 15 days to **fipronil** at DHI

The average body lengths of animals exposed to fipronil at DTU and DHI are shown in Figures 13 and 14. No effect was observed at DTU (up to 750 ng/L). At the highest concentrations of 1000 and 3000 ng/l at DHI a significantly reduced body length was observed for both males and females (Figure 14).

3.3.2 Egg Production

Figure 15 and 16 show the egg production at different times of exposure to 3,5-DCP and fipronil. An obvious problem with these observations is the low number of females present in each group making the statistics quite uncertain with large confidence intervals. Thus, it is not possible from these experiments to evaluate if egg production is a proper parameter to look at.

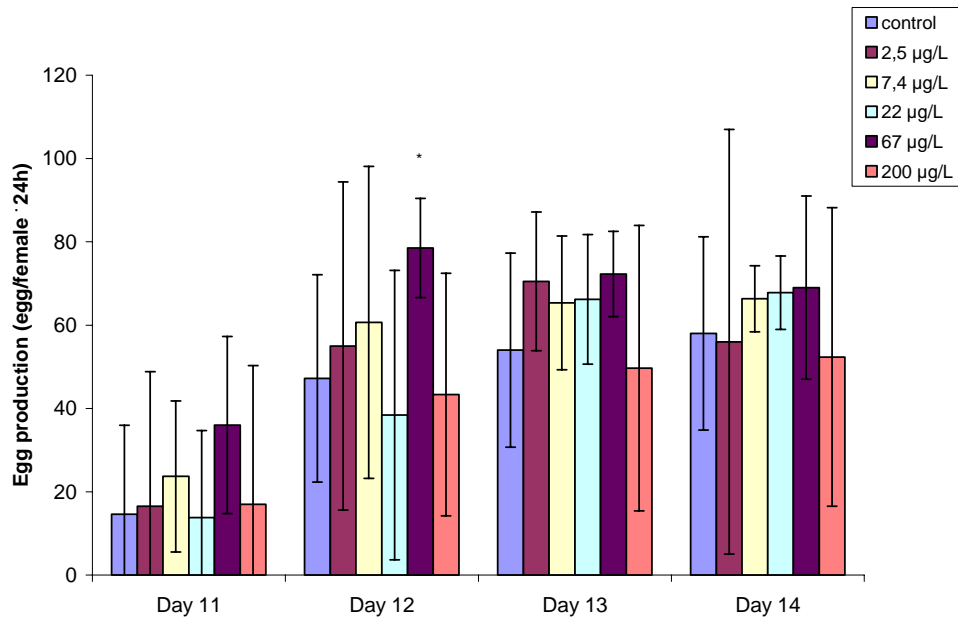
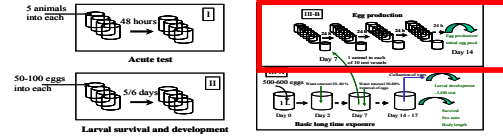


Figure 15. Daily egg production at different times of exposure to **3,5-DCP** at DTU. Asterisk shows significant difference from control (p-value < 0.05). Vertical bars show 95 % confidence limits of means. Only two to five females were present in each exposure group.

The onset of the egg production is difficult to evaluate. In most groups the egg production started on day 11. –At the highest concentration of fipronil (750 ng/L) egg production started already on day 10, but this group was only represented by one female, the remaining animals were males (6 animals) or died (3 animals).

The average egg production of the control females of the fipronil experiment was below the validity criteria. This is mainly due to a low egg production at the start of the reproductive period. At day 15, the egg production was well above 40 eggs per female and day in both tests. Temperature might also play a role. Unpublished data has shown that the development of *Acartia tonsa* is very sensitive to temperature changes. The allowed temperature variation of ± 2 °C thus seems to be too large. A narrower temperature range will also mean a more constant duration of the LDR test, where the exact time for termination actually is determined by the developmental stage of the animals.

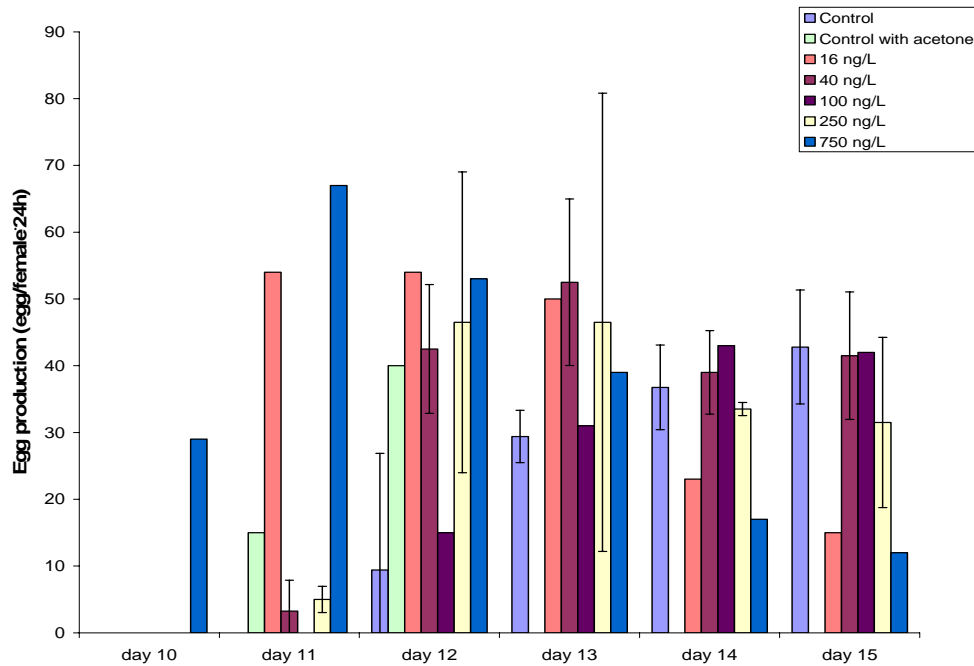


Figure 16. Daily egg production at different times of exposure to **Fipronil at DTU**. Vertical bars show 95 % confidence limits of means. Only one (without bars) to five females was present at each concentration.

An interesting observation was made in the egg production test with fipronil at DTU. The size of the eggs increased with increasing concentration. However, only at the highest concentration this effect was a statistically significant (see Figure 17).

The measurement of the egg diameter was initiated by the observation of some “double-eggs” at the highest concentration of 750 ng/l - eggs that either hadn’t separated as they normally should do or had fused during their development (see Figure 17).

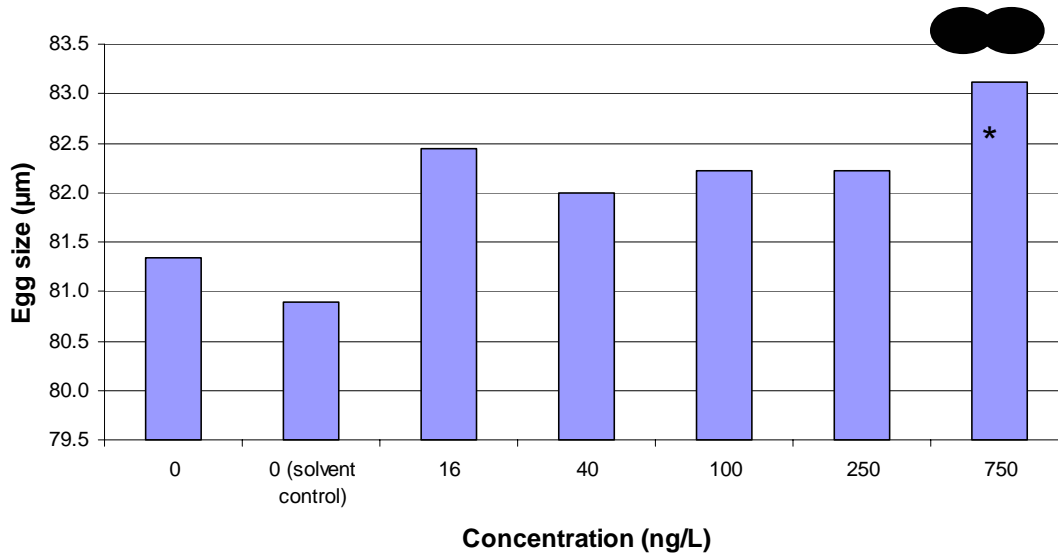


Figure 17. Average diameter of 25 eggs produced by fipronil-exposed females in the basic exposure at DTU. Asterisks indicate significant difference from control (p-value < 0.05). The black figure indicates the shape of a “double-egg”.

Figure 18 shows the egg production of *Acartia tonsa* exposed to fipronil at DHI. DHI investigated this parameter using a procedure deviating from the proposed standard method. At DHI the control consisted of 6 replicates with 3-5 females each and exposed groups of 4 replicates with 4-5 females. The females were isolated on day 14 and exposed for 7 days more. At day 18 and 21 after start of exposure, the number of eggs produced in each replicate was determined and the average number of eggs per female and day was calculated. The result is shown in Figure 18. One problem here is that the acetone control significantly differs from the control. Thus it is difficult to conclude on the effects of fipronil in this experiment

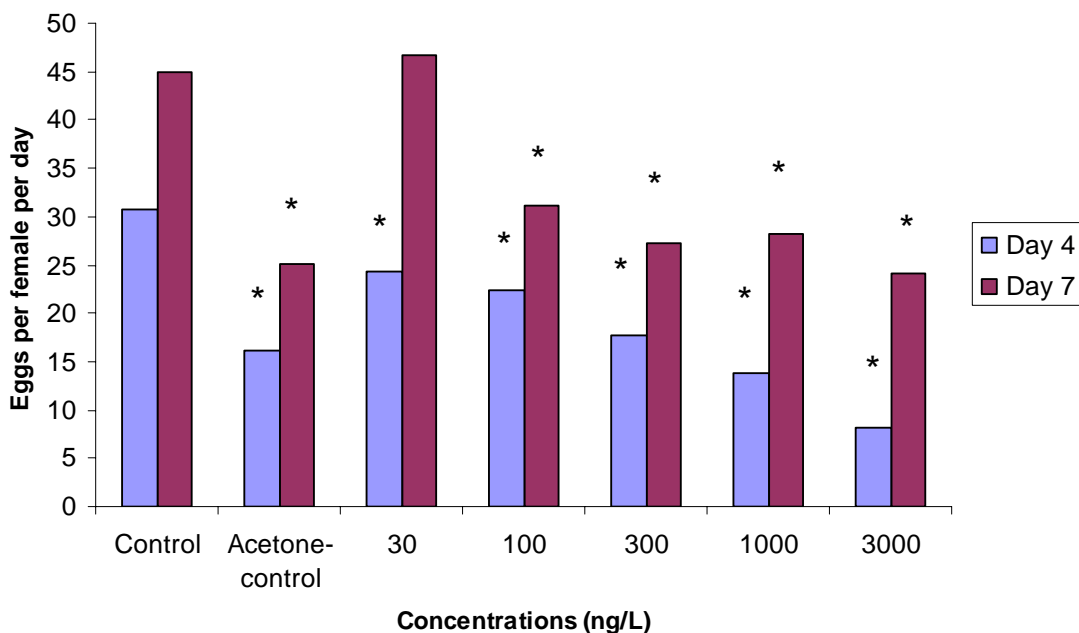
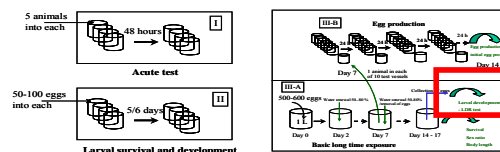


Figure 18. Egg production at day 18 (4 days after isolation) and day 21 (7 days after isolation of females) after start of exposure to **Fipronil at DHI**. The columns represent mean of 6 controls or 4 replicates with 4-5 females in each. Asterisks indicate significant difference from control (p-value < 0.05).

3.3.3 Second generation LDR

3,5 -DCP

The result of the second generation LDR test along with effects on hatching success and larval survival are shown in Figure 19. Like in the first generation test no effects were observed on hatching success or larval survival. Both parameters were high (> 90 %) in the controls as well as at all exposure concentrations and the validity criteria regarding hatching and survival were thus met. Significant effects on LDR were observed at 2.5, 67 and 200 µg/L of DCP. The concentration-response curve is shown in Figure 20 and the result of the log-normal statistical analysis is shown in Table 11.



From Table 11 it is seen that an EC10 and EC50 of 27 µg/l and 150 µg/L, respectively, were found. These values are a factor of 2-3 lower than those found in the first generation LDR (see Table 5).

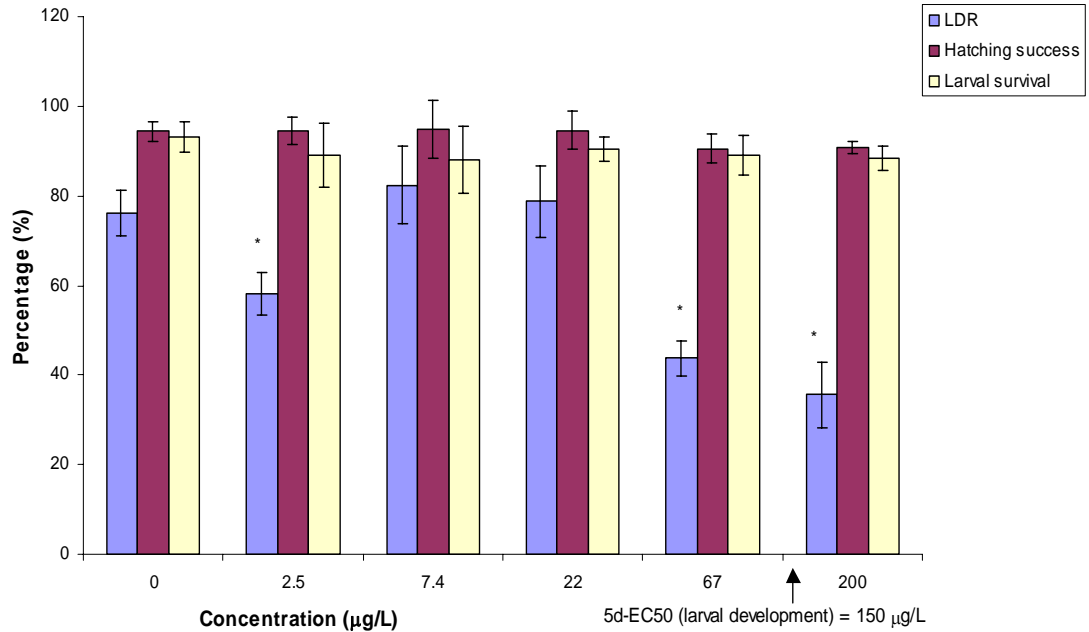


Figure 19. Second generation Larval Development Ratio (LDR = percentage of animals that reached a copepodite stage), hatching success and larval survival of *Acartia tonsa* exposed for 125 hours to **3,5-DCP** in the concentration range from 2.5 – 200 µg/L at **DTU**. Asterisks indicate significant difference from control (p-value < 0.05).

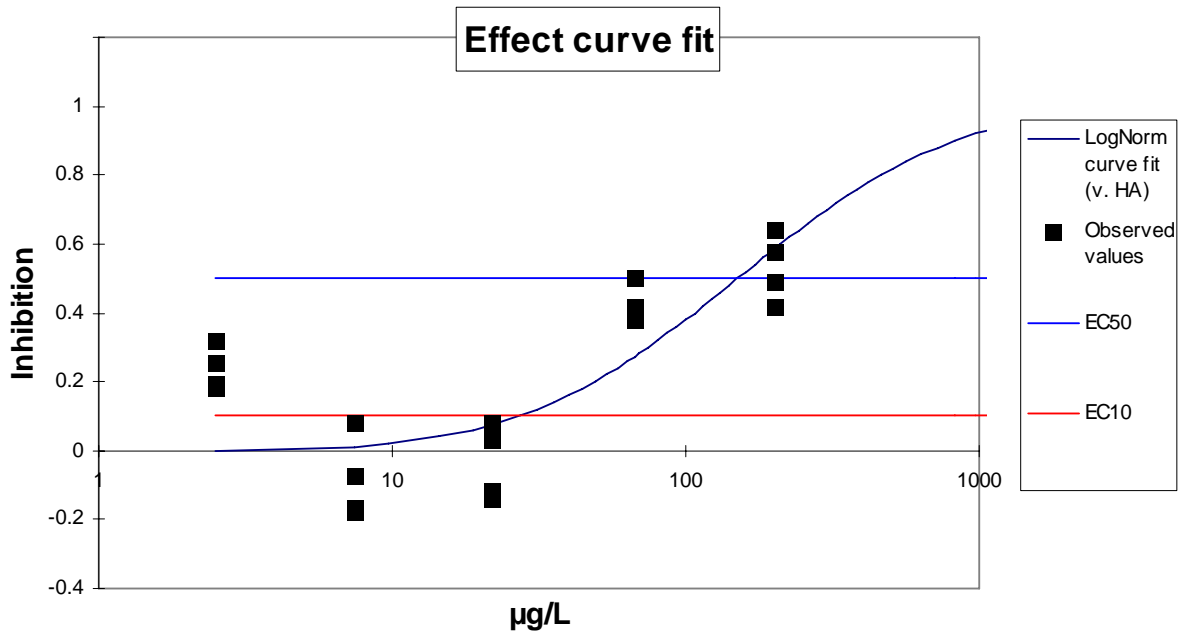


Figure 20. Concentration response curve for second generation LDR test with **3,5-DCP** and *Acartia tonsa* exposed for 139 hours at **DTU**.

Table 11. Calculated EC values for inhibition of LDR and confidence limits for the second generation LDR test with **3,5-DCP** and *Acartia tonsa* at **DTU**. Concentrations are given in $\mu\text{g/L}$ (nominal concentrations)

		95 % confidence limits	
EC	$\mu\text{g/L}$	Lower ($\mu\text{g/L}$)	Upper ($\mu\text{g/L}$)
10	27	12	61
50	150	91	248

Figure 21 shows the result of the second generation LDR test with **3,5-DCP** at **DHI**. Significant effects on hatching success and larval survival were found at the two highest concentrations of 500 and 750 $\mu\text{g/L}$. Like in the first generation test the control LDR was very low – not more than 20 % - and significant effects on this parameter were only found at 220 $\mu\text{g/L}$ and higher concentrations. The low LDR (with associated high variation - see Figure 22) was probably the reason why no significant effects were found at lower concentrations even though the mean LDR values were lower than 50 % of the mean control value.

Figure 22 shows the concentration response curve for the LDR test with 3,5-DCP at DHI and it is seen that most response values are above 50 % inhibition compared to the control. The calculated EC10 value of 10 $\mu\text{g/L}$ (Table 12) is very uncertain and outside the tested concentration range. The EC50 of 68 $\mu\text{g/L}$ is within the tested concentration range, but as in the first generation LDR test the copepodite fraction (LDR) in the control is low (~20%) and thus, the EC50 is uncertain. The EC50 value is approximately a factor of two below the one found in the parallel experiment at DTU (150 $\mu\text{g/L}$ – see Table 11).

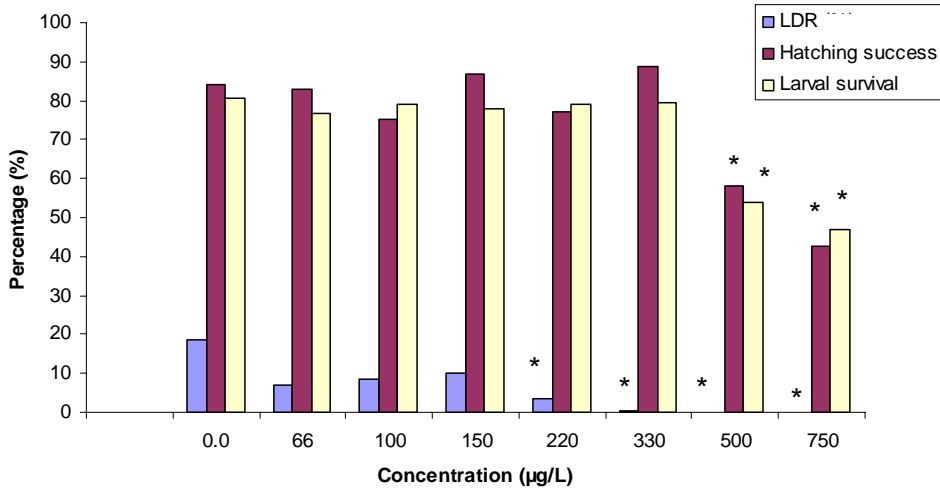


Figure 21. Second generation Larval Development Ratio (LDR = percentage of animals that reached a copepodite stage), hatching success and larval survival of *Acartia tonsa* exposed for 125 hours for 3,5-DCP in the concentration range from 66 – 750 µg/L at DHI. Asterisks indicate significant difference from control (p-value < 0.05).

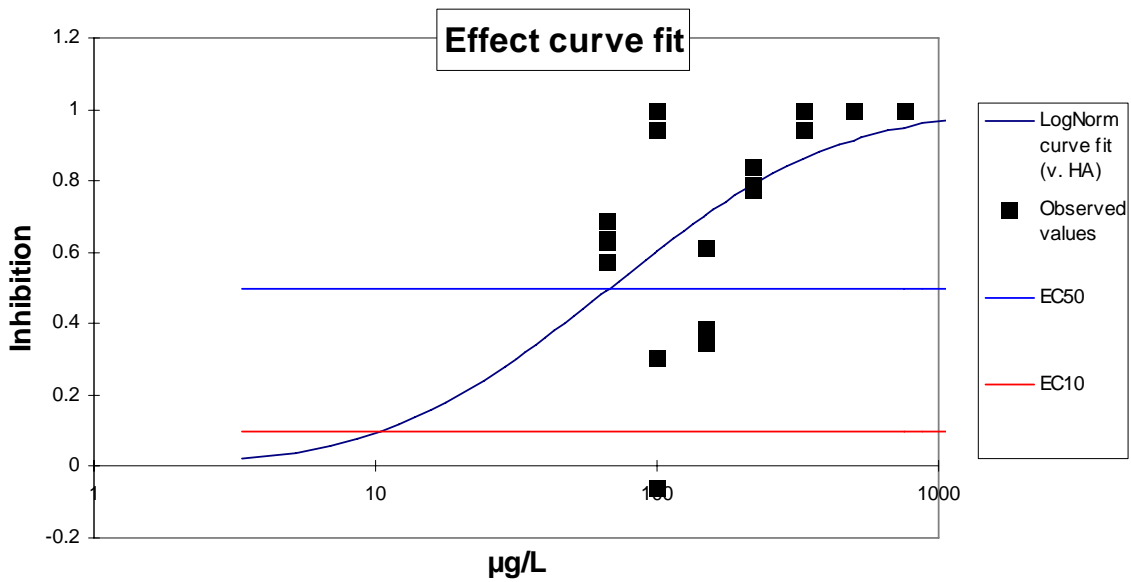


Figure 22. Concentration response curve for second generation LDR test with 3,5-DCP and *Acartia tonsa* at DHI.

Table 12. Calculated EC values with respective confidence limits for the second generation LDR test with **3,5-DCP** and *Acartia tonsa* at **DHI**. Concentrations are given in µg/L (nominal concentrations)

		95 % confidence limits	
EC	µg/L	Lower (µg/L)	Upper (µg/L)
10	10	1.5	70
50	68	32	145

Fipronil:

The result of the second generation LDR test with fipronil at DTU is shown in Figure 23. No effects of fipronil on larval survival were observed. At the two highest concentrations (250 and 750 ng/L) weak significant effects on hatching success were found. At 40 ng/L a significantly lowered LDR was found. The result of the second generation LDR test thus was close to that found in the first generation LDR test (Figure 6) and no concentration-response relationship was seen.

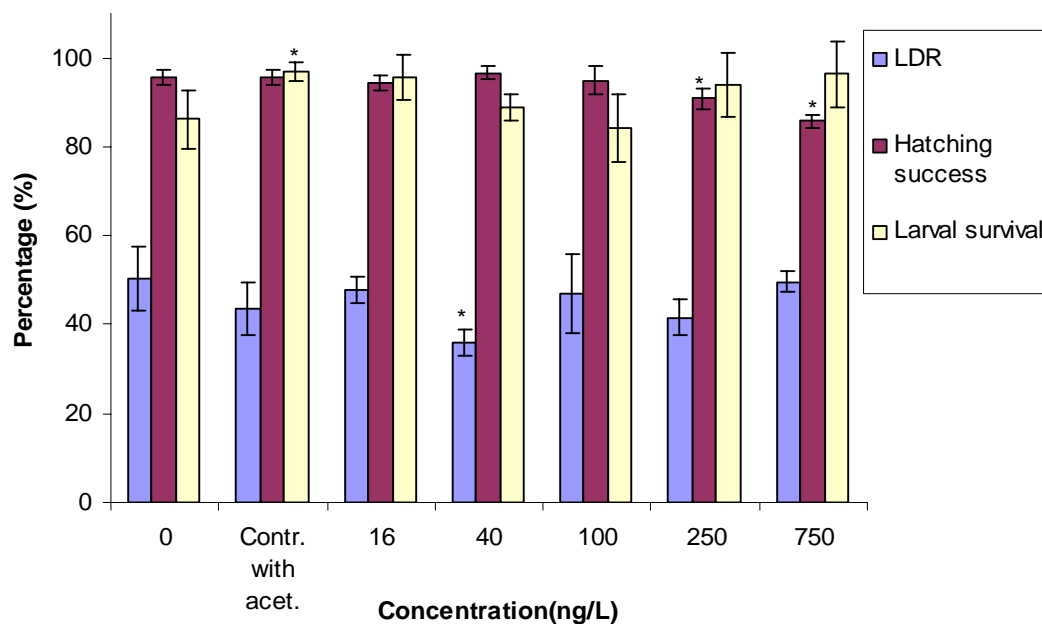


Figure 23. Second generation Larval Development Ratio (LDR), hatching success and larval survival of *Acartia tonsa* exposed to fipronil in the concentration range from 16 – 750 ng/L at **DTU**. Asterisks indicate significant difference from control (p-value < 0.05).

Figure 24 shows the DHI fipronil LDR test result. Significant effects on all three parameters were observed at the highest concentration of 3000 ng/L. The LDR in the acetone control was significantly different from that of the control, making it difficult to draw clear conclusions. If the result of the acetone control is ignored a concentration-response curve can be obtained (see Figure 25). The EC10 is extrapolated and even though the EC50 is within the tested concentration range there is a broad confidence interval for this value due to large variations in the replicate values at most exposure concentrations (Table 13). Thus, the EC50 value is uncertain due to problems with the acetone control and due to large variation among replicates.

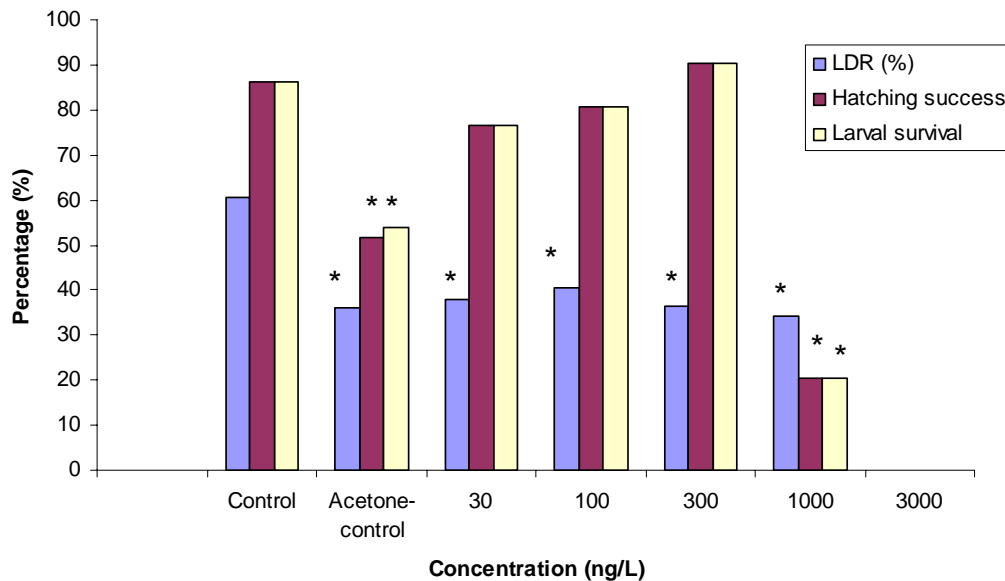


Figure 24. Second generation Larval Development Ratio (LDR), hatching success and larval survival of *Acartia tonsa* exposed to **fipronil** in the concentration range from 30 – 3000 ng/L at **DHI**. Asterisks indicate significant difference from control (p-value < 0.05).

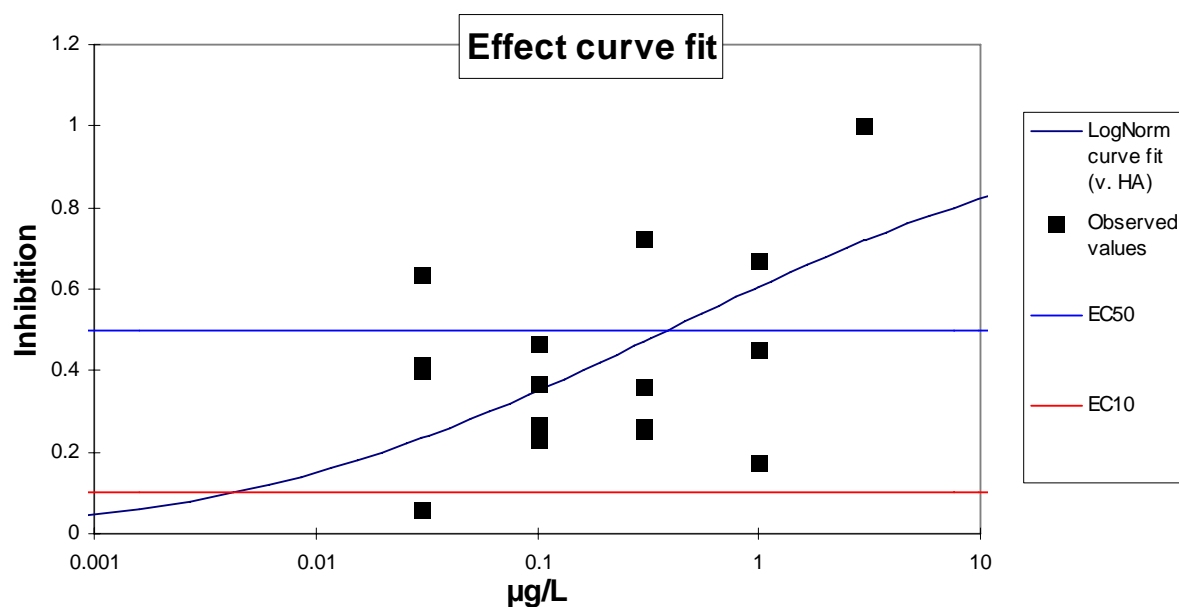


Figure 25 Concentration response curve for second generation LDR test with **fipronil** and *Acartia tonsa* at **DHI**.

Table 13 Calculated EC values and confidence limits for the second generation LDR test with **fipronil** and *Acartia tonsa* at **DHI**. Concentrations are given in µg/L (nominal concentrations).

		95 % confidence limits	
EC	µg/L	Lower	Upper
10	0.0042	0.000066	0.27
50	0.39	0.093	1.6

4 Chemical Analysis

4.1 3,5-Dichlorophenol

The analysis of 3,5-DCP failed probably due to an error during extraction procedure.

4.2 Fipronil

4.2.1 Stock solution

The stock solutions of fipronil were analysed at DTU using gas chromatograph with electron capture detector (GC-ECD) operated in splitless mode.

Three stock solutions were used and analyzed:

Table 14: Concentrations of fipronil stock solutions analysed by GC-ECD. (nominal concentration of 1000 µg/L).

Sampling date	Concentration in µg/L
2004-04-26	843
2004-05-11	725
2004-06-02	1056

A stock solution (2004-05-11) was sampled three times over a period of time corresponding to the test period and the result of the concentration analysis is shown in Figure 26. It is seen that the concentration of the fipronil stock solution was fairly constant around 600 µg/L.

Fipronil stock solution concentration

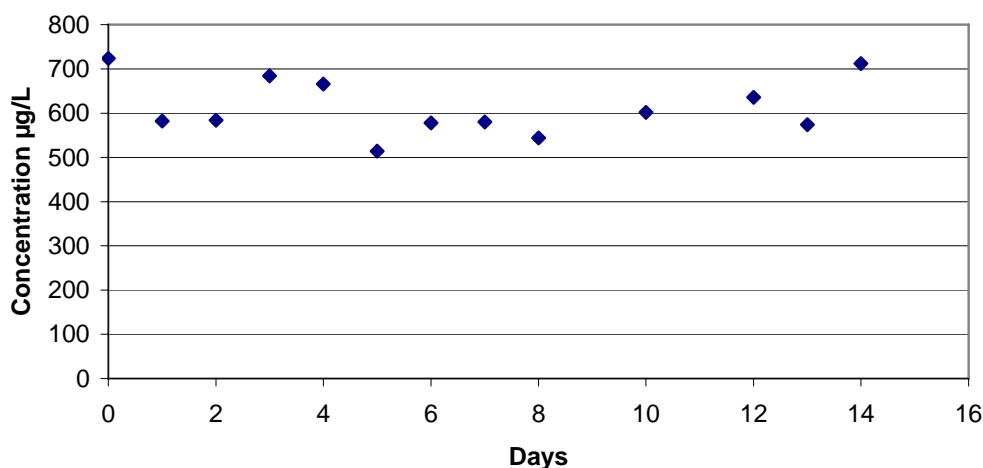


Figure 26. Fipronil stock solution (2004-05-11) analysed by GC-ECD over the test period

4.2.2 Water samples

Fipronil water samples were collected and stored at – 18 °C until analysis. Chemical analyses were performed by SOFIA GmbH, Berlin, Germany, using GC-MS. The limit of quantification was 30 ng/L.

Figure 27 shows the results of analysis of “used” media collected just before renewal of media. Analyses of water samples with three different nominal concentrations have been performed. Generally, the measured concentrations were about 50 % of the nominal in samples taken 2-3 days after renewal of the medium (see Figure 28). In Figure 28 are also shown the measured concentrations of two freshly prepared media samples, which amount to approximately 60 % of the nominal concentration. This corresponds to the percentage of nominal concentration measured in the stock solution.

Chemical analysis of fipronil

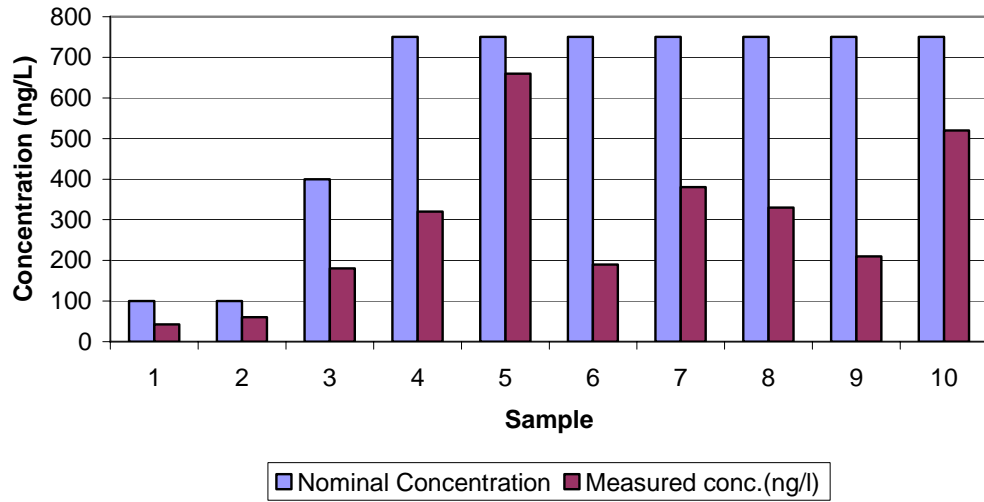


Figure 27. Nominal and analysed concentrations of fipronil in samples taken 2-3 days after renewal of the medium. Three nominal concentrations were selected for analyses.

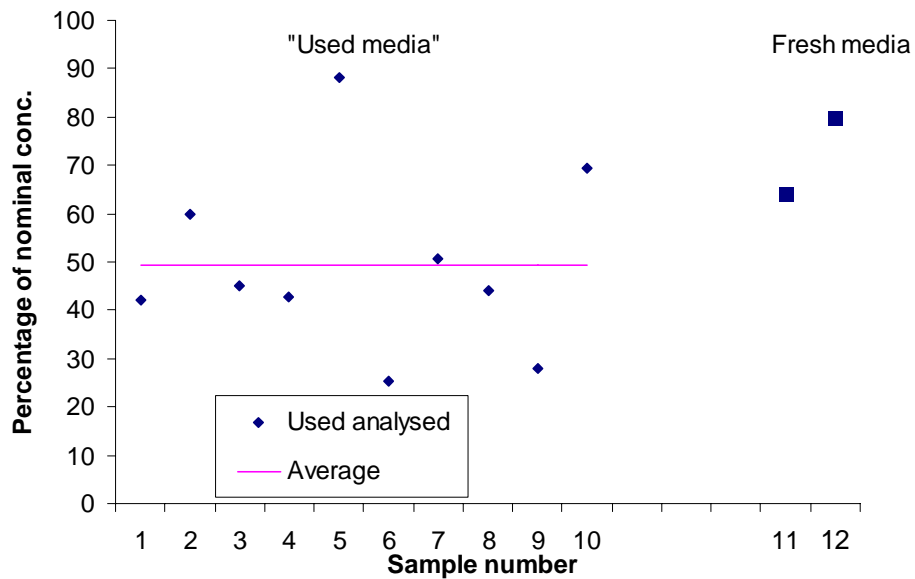


Figure 28: Percentage of nominal concentrations of fipronil in “used media” and freshly prepared media.

5 Discussion and conclusions

5.1 Chemical analysis

Since water samples from only one laboratory were analysed, nominal concentrations were used throughout the report in order to enable a comparison of results of the three laboratories.

5.2 Test results

5.2.1 Acute tests

The results of the acute tests conducted in the three laboratories were in the same range and for the reference compound 3,5-DCP the effect values met the performance criteria of the ISO standard.

5.2.2 Larval development ratio test on first generation (F0)

3,5-DCP:

Different ways of data treatment for the test with 3,5-DCP gave EC50 values of 192 µg/L (DTU) and 351 µg/L (DHI), respectively. Though not exactly the same EC50 were obtained, the tests at DTU and DHI gave comparable results.

Fipronil

Generally only weak effects were observed at the concentration ranges tested by both laboratories. At 10 µg/L DHI observed a strongly decrease in LDR along with reduced survival. This is not surprising, since 10 µg/L is in the acutely toxic concentration range (48 hr LC50 = 2.4-2.5 µg/L).

5.2.3 Life-cycle test

Hatching success:

The hatching control showed a high hatching success of more than 80 %. A similar high hatching rate was also observed in the F0 generation LDR tests referred to above.

Mortality in mass culture control:

The mortality in the control mass culture was as high as 57 % and generally more than 50 %. Thus, the validity criterion of 20 % control mortality seems not adequate. The 20 % was chosen as the same as in the *Daphnia magna* reproduction test, but since *Acartia* is a more short-lived organism it is not reasonable to use the same validity criteria for *Acartia*. Barata et al. (2002) also observed a lower survival of about 30-40 % after 14 days in a demographic effect study of cypermethrin. In a later study Medina and Barata (2004) showed a survival of 24-28 % after 14 days at individual densities of 500, 1000, and 2000 per liter. They concluded that at the highest density tested development was delayed and egg production was lower, but survival was not affected compared to the lower densities (Medina and Barata 2004). More experience is needed before a criterion on mortality can be specified.

Sex ratio:

The validity criterion of between 40 and 60 % of each sex in the control mass culture was not always fulfilled. In one case a control female ratio of only 37 % was observed and in several cases a ratio of close to 40 % was found. It may be questionable to keep this validity criterion since many factors influence the sex ratio of copepods. A closer study of this parameter and the factors influencing it needs to be performed.

Size of organisms:

DCP and fipronil were found to influence the body length of *Acartia* significantly. DTU found that DCP increased length of males and females at 200 µg/L and DHI found that 10 µg/l fipronil reduced the size of both sexes.

Egg production:

The DTU results regarding egg production of DCP and fipronil exposed *A. tonsa* were very uncertain due to an insufficient number of females present at some concentrations. When the animals are isolated their sex cannot be determined. Few females in a group results in high variation of the number of eggs produced per female. Thus, the proposed method is too uncertain to secure a proper number of egg producing animals. Also the onset of the egg production thus become a parameter that is only determined with a high uncertainty.

At the DHI groups of females were isolated at a later time of the test and the egg production was measured over 2 separate days. This modified method it is better – however the DHI experiment suffered from disturbances caused by acetone, which seemed to reduce the egg production – thus making it difficult to interpret the effect of fipronil on egg production.

A suggestion for an improvement of the egg production part of the draft guideline is to isolate 12 individual females at the time when the animals can be sexed – for example at day 14. Thus, a sufficient number of females can be secured reducing the uncertainty of the test result. The validity criterion for egg production then needs to be reformulated, but the number of 30 eggs produced per female and day can be kept.

In the fipronil test was observed an average egg production of the control females on day 12 below the validity criterion of 30 eggs • female⁻¹ • day⁻¹. This is probably due to a lower egg production at the start of the production period. Later in the experiment an egg production well above 40 eggs • female⁻¹ • day⁻¹ was observed. Since an influence of the temperature cannot be completely ruled out and since the development of *Acartia tonsa* is very sensitive to temperature variations, a modified criterion is suggested that the allowed temperature variation is restricted to ± 1 °C instead of ± 2 °C.

Larval development ratio test on second generation (F1):

The EC50 values found for 3,5-DCP in the second generation LDR test at DHI and DTU were 68 and 150 µg/l, respectively. Even though the DHI value of 68 µg/L is uncertain the values are in the same range. These values are about a factor of two-three below the EC50 values found in the first generation LDR test indicating an increased sensitivity of organism produced by animals, which were exposed during their entire life in the long-term test. The results of the two laboratories agree relatively well even though broad confidence limits were calculated.

A way of improving the results/reducing the uncertainties might be to use six exposed replicates instead of four in the LDR tests. The variability may be the same but the higher number of replicates will reduce the confidence limits. Also a recommended number of eggs of 50-80 (instead of 50–100) might improve the result.

Also a validity criterion for the copepodite fraction in the control at the end should be set up to secure a fraction of copepodites not too far away from the 50 % which is the optimal one. We will propose that the copepodite fraction shall be within 30-70 %.

Other parameters:

The observation of deformed eggs produced by fipronil-exposed animals led to the observation of a changed egg diameter. The egg diameter may thus be a valuable parameter to measure. This shows the importance of being aware of other effects than those described in the Guideline. Even though significant effect was only observed at the highest concentration the observation is interesting.

5.3 Suggested changes and overall evaluation

5.3.1 General changes

Reference compounds:

The strong acute toxicity of fipronil caused problems since it seems to be the dominating effect of this compound. Furthermore, the lowest concentrations of fipronil used in the life-cycle test are below the quantification limit of the chemical analysis.

Unfortunately, the 3,5-DCP cannot be expected to exert any endocrine disrupting effects. Parameters reflecting such effects of fipronil cannot be evaluated since they are overridden by the high acute toxicity of that compound. .

Thus it is recommended to replace fipronil by a more suitable reference compound. Bisphenol A is suggested as new reference compound. It has shown estrogenic effects in test systems with various organism groups and there are indications that it interferes with the hormonal systems of calanoids as well (see Annex 1).

Inclusion of the F0 generation in the guideline:

Based on the incoming comments it has been suggested to include the LDR test on the F0 generation in the test guideline. In the draft prevalidated it was optional to perform the F0 LDR test. The test gives a good base for selection of concentrations for the life-cycle exposure and - when results are compared with those from the F1 generation - important information on changes in effects on two different generations is gained. Besides, it is not labour intensive to perform.

Temperature variations:

The development of *Acartia tonsa* is very sensitive to temperature variations and a modified allowed temperature variation is suggested that restricts the variation to ± 1 °C instead of ± 2 °C.

Repeatability and accuracy:

Data for evaluation of repeatability and accuracy are limited. Only for the LDR test on the first and second generation with 3,5-DCP there are two data sets to compare and these gave EC50 values that only differ by a factor of two to three and thus are in the same concentration range.

An increase in the number of replicates in the LDR tests from 4 to 6 can improve the accuracy of the results and reduce the confidence limits and will be proposed as an option in the new draft guideline.

Egg production:

In the draft prevalidated egg production of animals isolated on day 7 of the basic exposure was measured. At this point of time the sex can not be determined. In some cases this resulted in very low numbers of females and corresponding insufficient data material for statistical analysis. DHI used a different method by isolating females at a later point of time. In previous studies DTU measured egg production of groups of 2 weeks old females. At this stage the egg production is stable and the number of eggs produced is well above the criterion value of 30 eggs•female⁻¹•day⁻¹ (see Table 2 in Annex 1).

A suggestion for an improvement of the egg production part of the draft guideline is therefore to isolate 12 individual females at a time when they can be sexed – for example after 14 days. Thus, a sufficient number of females are secured, thus reducing the uncertainty of the test result. The number of eggs produced within 24 hours is counted at least three days (on day 2, 3, and 4 after isolation). At the first day after isolation egg production is often significantly lower than the following days and should therefore not be included to the data set. The validity criterion needs to be reformulated slightly:

- the average control egg production shall be higher than 30 eggs•female⁻¹•day⁻¹ each of the three observation days from day 15 to day 18

5.3.2 Changes of validity criteria

Larval development ratio:

In the draft it was recommended to stop the LDR test when the copepodite fraction is as close to 50 % as possible, but no range was given. The best chance to observe a slower or faster development of exposed groups compared to the control is at the point where the control copepodite fraction is 50 %. It is not practically possible to stop at exactly this point – some deviation from the 50 % will normally be seen, but there should be a limit for the variation to ensure that the test has not run too long or too short. Therefore, a validity criterion of 50 % ± 20 % is suggested for the control copepodite fraction at the end of the F0 as well as F1 generation LDR test.

Survival in the LDR test:

The criterion for survival of control animals (≤ 20 %) in the LDR test was in some cases not fulfilled. An analysis of previous data revealed that a control survival of 70-80 % frequently was found, though it generally was higher (see Table 1 in Annex 1).

With an allowed mortality of up to 30 % most of our previously performed tests would have passed this criterion. It is therefore suggested to change this criterion to:

- average mortality in the controls on the day of observation of LDR should not exceed 30 % of hatched animals.

Survival in the control mass culture:

The present demand regarding survival in the control mass culture is clearly too strict (< 20 % mortality) since *Acartia* is a short-lived organism. Natural (background) mortality alone will cause the mortality to be much higher. In the present studies, between 50 and 60 % mortality was found. This is in line with the mortality found in a study by Barata et al. (2002) who observed a control survival of about 30-40 % after 14 days in a demographic effect study. This relative high natural mortality of *A. tonsa* is related to the high reproductive output. For comparison, in the harpacticoid test with *Amphiascus tenuiremis* 18 nauplii were produced over 12 days (two broods) (Chandler 2005) whereas *Acartia tonsa* produces about 50 egg • female⁻¹ • day⁻¹. This has to result in a shorter life span and a higher natural background mortality for *Acartia tonsa* compared to *Amphiascus tenuiremis*.

More experience is needed before a validity criterion for control survival in the mass culture can be formulated. It should be noted that mortality in the mass culture was intended to be only a performance parameter and not a test parameter.

Sex ratio in control mass culture:

The demand regarding the sex ratio also needs revision. Many factors influence the sex ratio and the present criterion of between 40-60 % of each sex may thus often not be fulfilled. It is suggested to leave out this criterion in the guideline until more experience is gathered about the factors influencing sex ratio in *Acartia tonsa* cultures.

On the other hand, (see comments by H. Tyle in Annex 1) the sex ratio and the control mortality are both relevant as performance criteria. At the end of the basic exposure there needs to be enough animals to perform the egg production test as well as animals for measuring of body lengths.

A new criterion therefore is suggested:

- At the end of the basic exposure at least 50 females and 50 males shall be alive in the control.

This new criterion will ensure that there are animals for both measuring of body lengths and for the egg production test. Besides, a lot of work is saved because the exact number of egg at the start of the test doesn't need to be counted, and sex determination of all animals in the control at the end of the basic exposure is no longer necessary (only 100 animals need to be sexed). If the survival is very low and the sex ratio is very far from 50:50 this criterion may not be fulfilled, indicating insufficient experimental conditions.

5.3.3 General conclusion

Based on the gained experience from the pre-validation work it is concluded that a revised draft guideline can be worked out and be ready for ring-testing.

References

Andersen, H. 1994. Statistical methods for evaluation of the toxicity of polluted water. Master's thesis at IMM, Technical University of Denmark, Kgs. Lyngby, Denmark. (In Danish)

Andersen, H.R., Wollenberger, L, Halling-Sørensen, B and Kusk KO. 2001. Development of copepod nauplii to copepodites - A parameter for chronic toxicity including endocrine disruption. *Environ Toxicol Chem* 20 (12): 2821-2829.

Barata, C., Medina, M., Telfer, T. and Baird, D.J. 2002. Determining demographic effects of cypermethrin in the marine copepod *Acartia tonsa*: Stage-specific short tests versus life-table tests. *Arch Environ Contam Toxicol* 43: 373-378.

Bjørnstad, E., Petersen, G.I., Robson, M., Reiersen, L.-O., Henriquez, L., Massie, L. and Blackman, R. 1993. Paris Commission ring test: testing of offshore chemicals and drilling mud on selected marine organisms. *Science Total Environ Suppl* 1993: 713-719.

Cary, T.L.; Chandler, G.T.; Volz, D.C.; Walse, S.S.; Ferry, J.L. 2004. Phenylpyrazole insecticide fipronil induces male infertility in the estuarine meiobenthic crustacean *Amphiascus tenuiremis* *Environmental Science & Technology* 38 (2): 522-528.

Chandler, G.T. 2005. Pre-validation of a copepod full lifecycle bioassay with the harpacticoid copepod *Amphiascus tenuiremis*. Dept. of Environmental Health Sciences, Arnold School of Public Health, University of South Carolina, Columbia, SC 29208. Spring 2005.

Connelly P. 2001. Environmental fate of fipronil. California Environmental Protection Agency. www.pw.ucr.edu/textfiles/fipronil.pdf.

Huyskens, P., Mullens, J. and Gomez, A. 1975. Solubility of alcohols, phenols and anilines in water. *Bull Soc Chimiq Belges* 84 (3): 253-262

International Organization for Standardization 1998. Water Quality - Determination of acute lethal toxicity to marine copepods (*Copepoda, Crustacea*) ISO/FDIS 14669.

Kusk, K.O. and Wollenberger, L. 1999. Fully defined salt water medium for cultivation of and toxicity testing with the marine copepod *Acartia tonsa*. *Environ. Toxicol. Chem.* 18 (7): 1564-1567.

Mauchline J 1998. The Biology of Calanoid Copepods. *Advance in Marine Biology* Vol. 33. Academic Press, San Diego.

Medina, M. and Barata, C. 2004. Static-renewal culture of *Acartia tonsa* (Copepoda: Calanoida) for ecotoxicological testing. *Aquaculture* 229: 203-213.

OECD 2003. Final Report on First meeting of the ad hoc expert group on invertebrate testing. 23rd – 24th November 2003 – OECD, PARIS. 2003.11.28.

OECD 2004. (Final) Draft Document OECD Draft Guidelines for Testing of Chemicals. Proposal for a new Guideline. Calanoid Copepod Development and Reproduction Test with *Acartia tonsa*. March 2004.
Test Guideline prepared by K. Ole Kusk and Leah Wollenberger for OECD Paris.

Shiu, W.Y., Ma K.C., and Varhanickova, D. 1994. Chlorophenols and alkylphenols - a review and correlation of environmentally relevant properties and fate in an evaluative environment. CHEMOSPHERE 29 (6): 1155-1224.

SPARC ON-LINE CALCULATOR – for calculation of pKa values.
<http://ibmlc2.chem.uga.edu/sparc>

U.S. Environmental Protection Agency. Ecotox database.
http://www.epa.gov/cgi-bin/ecotox_quick_search

Zarorc-Končan, J., Gotvajn, A Z and Tisler, T. 2002. Hazard identification for 3,5-dichlorophenol in the aquatic environment. Cellular & Microbiol Lett 7 (2): 381-382.

6 Annex 1 – Comments

Comments on draft report by Henrik Tyle Danish Environmental Protection Agency

Dear all.

I have read the two reports in a plane the other day on one of my many trips. Here are as promised my comments from the notes I made:

- 1) Many technical improvements of the methods have been elucidated during these pre-validation exercises.
- 2) Both reports are draft interim reports I assume (because the results from collaborating labs are not yet received/ reported: when is the deadline for finalisation of the reporting which could be to give up waiting for input ?) I would like to know when the final reports are to be expected.
- 3) It is recommended to consider harmonising the chosen statistical analysis between the two reports (DTU perhaps considering using ANOVA ?, ITM to specify curve fitting function. Shouldn't the curve fitting function be justified ? the EC10 could be heavily influenced by the curve fitting function ? The results of the follow up ring-test should be done by optimal statistical analysis -- and done the same way for all labs (if results from single labs are to be reported)
- 4) note that the analytical measurement of DCP failed (DTU) was not performed? (IMT) (consider impact of allowed pH variation (pH var. Of 0.3) on the conc. of unionized DCP in the test media.) Assume that analytical measurement of DCP in place and to be included in follow-up ringtest (validation)?
- 5) note fibronil recommended by IMT bases on some preliminary conclusions (not so clear to me but do not relate to conc.-effect curves obtained ?) - note that fipronil Not recommended by DTU - so what is overall recommended? The explanation regarding fipronil is not easy to follow - and explanation not similar in the two reports - consider using the same and clear explanation. Assume that the second ref. chemical for the ring test should be the same for the two methods.. DTU proposes BAP (please see the EU RAR on the ECB website with agreed relevant data & evaluation concerning relevant intrinsic properties incl. toxicity to crustacean sp. Consider analytical procedure for BAP incl. prize.)
- 6) purity of test chemicals not specified -. Should be in ring test - also specify if all labs are using the same batch (distributed by lead lab)
- 7) Consider incl. acute test and F(0) part in the TG because these referred to as "essential" for optimal selection of test conc.s Incl. recommendation on how to make this choice optimal (as regards spacing of the test concs in the pre-validation a factor of approx 3 between test concs but no real explanation for the test conc. interval presented)

8) Consider harmonising as much as possible the performance criteria. Here the A. tonsa method seems to have dropped some of the originally set criteria (e.g. on control survival and sex ratio - but is these factors not relevant as performance criteria - what are the implications of simply dropping them ? Should more pre-validation to find out be considered before ring-testing ?) and the knowledge base for this test is perhaps less than for N.spinipes ?

Best regards

Henrik Tyle

Miljøstyrelsen

Response from Ole Kusk, DTU, to Comments by Henrik Tyle, Danish Environmental Protection Agency- and Henrik Tyle's counter-response

Dear Henrik

I just spoke to Estelle Bjørnstad at DHI about your comments and we agreed to the following responses to your comments (for the Calanoid part of the project):

1) -

2) I have got comments only from you. Tim Williams has announced that he will give some comments as well. I will wait for further comments to come in until mid August end then finalise the report.

Henrik: good idea to fix end Aug for finalisation of input to the reports

3) The curve fitting function is described at page 10 in the calanoid report. It is a fit to the log-norm function. We have also the possibility to fit to the Weibul function in our program, and there is a difference between the two and mostly on the EC10 value, but generally we find only minor difference between the two types of fittings. Also, in the ring-test we shall use the same type of statistics (ANOVA is the used more than t-test I think, but both have their advantages and dis-advantages).

Henrik: please note that I asked for harmonisation between the two tests and argumentation / justification for the selected methods. I asked about the curve fitting procedure. Nice to know that these two functions does not really impact the EC20 - but how well did they fit?

4) I think we all have a problem with the financing of the chemical analyses in the ring-test. Not sure what to do about it.

Henrik: Something HAS to be done!: has to be solved and done, because all aquatic (pelagic) OECD TGs require chemical analysis. (even the algae test where it can be somehow problematic to prove maintenance of water conc. by comparison of initial and end conc (at least if measured on some hydrophobic chemicals and by measuring in the

culturing medium with living algae..) This problem could potentially be a severe one: I suggest that both drafters
- NU steering gr. And their authorities consider this further - more input regarding 2. ring test chemical, costs etc needed from the two drafters

5) We can not use fipronil as a reference compound for Acartia. There are then (at least) the following possibilities:

- use of only one reference compound which then should be 3,5-dichlorophenol (3,5-DCP)
- use of 3,5-DCP with both copepod groups plus fipronil for harpacticoids and Bisphenol A for calanoids
- use of 3,5-DCP and a new reference compound for both types of copepods (but which one??? - we need a very quick decision in this case)

Henrik: To me the more obvious choice would be for both as the second ref. chemicals to use bisphenol A - what about analysis cost etc. (cf. the EU RAR on ECB's web site - possibly if needed I can help you finding more and possibly not yet published info on this substance where a further research follow up program is running under the ESR programme.

6) More info on test chemicals (purity, batch no, etc) will be included in the final report. I think the harpacticoid group used common batches, and we in the calanoid group also used common batches, but different from the ones used in the harpacticoid group.

Henrik: *OK - in the ring test preferably all should use specified chemical from the same batch.*

7) It might be a good idea to include the LDR on F0 in the TG, but not the acute tests. There is already an ISO standard on the acute test, and the present guideline is meant for investigation of chronic effects. Including the F0 LDR will give a good base for selection of proper test concentrations in the life-cycle test, and it will also give the base for evaluating if there's a greater effect at the second (F1) generation.

Henrik: *Based on your response I guess the LDR on F0 should be included in both TGs...*

8) The sex ratio varies from batch to batch of Acartia and thus I think it's problematic to have the sex ratio as a validity criterion. We don't know all the factors that control this ratio. The other criteria we dropped was the control survival and it was somewhat higher than we expected (between 50 and 60 %) and this was confirmed by the literature. We could of course change the criteria to less than 70 % control mortality to secure a sufficient egg production, but if in a test there are too few eggs due to control mortality there is a big problem with this test and it should simply be repeated.

And I don't think we need more prevalidation.

Henrik: *Note that you don't think more pre-validation is required. So:*

Sex ratio: do you think there is "a bottom line" (min. acceptable

performance) under which you feel that a repeat study with another and more balanced sex ration is required. If so that should be set as a performance criterion (e.g. 20:80 or vise versa - or whatever ration you think is needed, would be nice form your's and others experience to know more about the observed variation) Control mortality: Comment similar to that concerning sex ration: what is thgen the suggested max parental control mortality: 50 or 60 % - how by the way is reproduction calculated: per surviving mother (or is it per parentanimal?), per initial mother ?? (similar discussion has taken place concerning the daphnia magna test)

I hope this answer the questions you put up in your comments

Best regards

Ole

K. Ole Kusk

Comments on draft report from Tim William, AstraZeneca UK Limited:

From: Tim Williams
(contact details below)

Date: 15 July 2005

COMMENTS CONCERNING:

Report to Nordic Council of Ministers and OECD – Validation of full life-cycle test with the copepod *Acartia tonsa*

General comments:

The authors describe the *Acartia* validation work for two compounds at two different laboratories. The results demonstrate that good progress has been made and this report provides a useful document for making informed decisions on the status of the *Acartia* life-cycle test as an OECD guideline.

The data presented suggest that further validation work is required before this toxicity test could proceed in a ring test. Key parameters such as larval development ratio and egg production failed to meet the 'original' validity criteria of the test guideline and a number of major revisions to the draft guideline have been proposed as a result of this work. Most importantly the reproduction phase of the life-cycle test has been completely redesigned but as yet there is no reliable data to suggest that this new design will provide repeatable data from which toxic effects of a test compound can be established.

There is some debate over the most reliable endpoints to measure in this life-cycle test. For example, unless the increase in egg size observed in the fipronil study was a common toxicological response it would probably be a time consuming endpoint to measure. This type of response would probably be detected as part of routine morphological and behavioural observations.

Larval development ratio appears to be one of the most reliable test endpoints but both laboratories could not consistently meet the validity criterion. The authors suggest temperature as a possible reason for low control reproduction in one of the studies. More strict temperature control may also reduce the variability in the larval development ratio.

A validity criterion of 50-60% mortality seems very high even given the argument that *Acartia* is a short-lived species. Even in mollusc and echinoderm embryo larval tests mortality is expected to be less than 30%. It is important to show that the culture and handling conditions of the copepods are optimal. A number of parameters measured in these studies suggest that the culture conditions for *Acartia* may be sub-optimal including differences in the larval development ratio between laboratories, high mortality and variable reproduction.

Recommendations

Considering the proposed changes to the Guideline following this pre-validation work, it may be advisable to proceed with further experimental work (e.g. repeat the DCP exposure experiment based on the updated draft guideline) before ring testing. This would establish that an experienced laboratory could achieve the new validity criteria and that the reproduction phase of the test could provide reliable data.

If the validity criteria for mortality were to be higher than 30% there would need to be greater justification for this. The authors need to provide stronger evidence that the observed mortality is a feature of the biology of the test species and not due to sub-optimal culture and handling conditions.

Specific comments:

Preliminary tests measuring acute toxicity and Larval Development Ratio (LDR)

Refer to pages 10-11 (acute toxicity) and 12-19 (LDR).

There appears to be relatively good agreement between the 3 laboratories providing LC50 data. Comparing 48 h LC50 values, the results for DHI were slightly higher than those of DTU & ECT, however, DHI were using natural seawater as opposed to synthetic seawater. Results may be influenced by subtle differences in the water chemistry between the two sources (natural versus synthetic).

For the preliminary LDR test on the F0 generation, both laboratories (DTU & DHI) appeared to achieve good egg hatching success, however, larval survival was sub optimal in the DTU test (approx. 70%) and failed to meet the Guideline criterion of >80% survival. The issue of natural versus synthetic seawater is an interesting one! The LDR test results for 3,5-DCP and fipronil from DHI, which used natural seawater, generally achieved >80% survival (met guideline criterion) whereas those for DTU (using synthetic seawater) did not. Is the difference in larval survival due to the choice of test medium (i.e. natural seawater versus synthetic seawater!)? Note,

however, that in the second generation LDR test (see below), the results for larval survival from DTU were >80%.

Concerning the measurement of LDR by DHI, larval development was consistently low (<30%) in control treatments although no explanation was given. DHI stopped their LDR tests after exactly 120 h (5 days), which under the test conditions, may have been too early (5-7 days in Guideline). If poor larval development was apparent after 5 d, why not simply perform the measurement of LDR later (i.e. after 7 d), to allow sufficient time for an acceptable number of nauplii to develop into the copepodid stages? No data were provided for physical parameters during the test but could delays in development be due to variation in test temperatures? Based on practical experience, did DHI expect to achieve at least 50% LDR in controls after only 5 days?

Although this is a preliminary test, the comments are relevant for the life-cycle test as LDR is measured in the F1 generation. As a general comment, why does the method not aim for a higher portion (e.g. 80% as opposed to 50%) of nauplii reaching the copepodid stage? As the copepods are retained for approx. 14-17 days, LDR could be measured later than 5-7 d, when higher proportions of nauplii have attained the copepodid stage. When subsequently comparing the results of LDR in control versus treatments, it may be easier to detect delays in development based on a higher LDR in the control!

Second generation Larval Development Ratio (LDR):

Refer to pages 27-33.

The results from DTU for the second generation LDR test (see Figures 19 & 23) met the validity criteria, however the LDR results for DHI (see Figures 21 & 24) were, like the preliminary LDR test, very low; LDR results for fipronil [>50%] were the exception, however, these results also indicated some negative effects of the carrier solvent. It would be very helpful if the authors can explain the reasons for such differences (e.g. was adequate temperature control provided in the DHI experiment?). By comparison with egg production, LDR in *Acartia* would appear to be an easier parameter to measure. However, as with egg production, it is important to explain/resolve the differences in LDR between the two laboratories (DTU & DHI). DHI carried out some amendments to the protocol (e.g. use of natural seawater), however, no discussion was included as to the influence of these changes on the test results.

Discussion on LDR

Response to points made in discussion of test results (refer to pp. 36-38).

The authors recommend increasing replication as a means of improving the results. The replication for the second generation LDR test is not clear – the Guideline states that 50-80 eggs are exposed in each replicate but the numbers of replicates per treatment do not appear to have been stated, however, it is assumed to be the same as for the preliminary LDR test (4 replicates per treatment). Concerning within-treatment variability of the LDR data the results provided by DTU (Figures 2, 7, 19 &

23) do not indicate what the error bars denote (standard error mean or standard deviation). The results for DHI (Figures 4, 8, 21 & 24) also omit this data, probably because of the very low LDR values (with the exception of the fipronil data shown in Figure 24), therefore, it is difficult to comment on the issue of variability in the LDR data.

The results for LDR suggest that larval development of *Acartia* is a sensitive and valuable technique for measuring effects of toxicants. However, variability between laboratories was high (DTU >50%, DHI generally <30%). I am not sure that the factors responsible for these differences are well understood (e.g. inadequate control of temperature), and further experimental work may be needed to do so. The authors propose to include a validity criterion for LDR makes good sense, however, further discussion on what this value should be would be welcome (see below).

The authors recommend a validity criterion for LDR, such that the copepodid fraction in controls at the end of the test should be within 30-70% (20% either side of the 50% stated to be 'ideal'). As discussed earlier, whilst LDR can be measured at different time periods, you are looking to 'capture' the transition between nauplius to first copepodid stage. Consequently, I do not understand why the exposure period (e.g. 7 days at a temperature of 20°C) cannot be selected to ensure a relatively high proportion of animals in the copepodid stage (e.g. >70%). The results presented by DHI did not capture the transition from nauplius to copepodid stage (LDR was very low) therefore, simply extending the range of the validity criterion to accommodate poor larval development (in the time period specified) seems to ignore the key issue of why a significant proportion of nauplii do not reach the copepodid stage as expected, within the specified time period.

LIFE-CYCLE TEST (refer to pages 20-33)

Mortality in mass culture and sex ratio

Refer to pages 20-22.

Considering the results for DTU, complementary egg hatch in control medium met the validity criterion (>80%) indicating that the batch of eggs was acceptable for the test. However, in both the 3,5-DCP and fipronil tests, control survival (after 14 d) was consistently low (52-57%), well below the validity criterion of 20%.

As a point of interest, clarification is needed here that mortality includes eggs that fail to hatch (e.g. 20 % failed to hatch in the complementary control hatchability test) and larval mortality after hatching. For *Acartia*, control hatchability tests generally reveal that not all eggs hatch (approx. 20% fail to hatch). In calanoid copepods, the concept of 'resting eggs' is well known although not particularly well understood. However, it is feasible that under the culture/test conditions employed in the laboratory, a proportion of the eggs produced are resting (i.e. dormant) eggs and would not be expected to hatch. If production of resting eggs is an inherent feature of the biology of this species, then failure of some eggs to hatch may not necessarily be due to the test chemical.

Regardless of whether mortality includes egg hatch and larval survival, mortality after 14 days is very high. No explanation is given, however, these results suggest a fundamental problem with the culture/test techniques, perhaps due to due aspects of the biology of *Acartia* (e.g. water quality requirements, food quality and quantity, handling techniques etc). Comparable results are not available for DHI, however, are mortalities of this magnitude typical for this species? Such high ‘background’ mortality, if an indication of poor health or suboptimal conditions, may raise concern about the validity of the test results (false positives).

The sex ratio during the F0 stage of the life-cycle test sometimes departed from 40-60% females. Similar results are sometimes observed in harpacticoid copepods, for example, sex ratio in *Tisbe* spp. is particularly sensitive to factors such as culture density. In the *Acartia* test, the bias in sex ratio towards males has important implications for the measurement of egg production in animals (n=10 per treatment) isolated from the mass culture initiated with 500-600 eggs on exposure day 0. It would appear that DTU isolate animals for the egg production experiment on exposure day 7, which is surprising given that these animals are not likely to be larval stages and not easy to sex. If it is not possible to confidently sex these animals at the selection stage (i.e. day 7), the observed sex ratio of approximately 40% females predicts that only 4 of the 10 animals randomly selected would be female. For a robust statistical analysis of the egg production data, taking into account that variation in egg production between individual females within a treatment may be high, I would question whether 4 animals are sufficient. Presumably, this explains why DHI chose to select animals at day 14, because you can more easily identify the sex of each animal thereby ensuring that you are able to select at least 10 females per treatment for the egg production phase. Compared with early selection (e.g. day 7), selection at a later time (e.g. day 14) may also ensure that all females have been successfully fertilized. The egg production data shown by DHI (Figure 18) appear to be just within the Guideline criterion (average number of eggs per female per day in the controls should not be lower than 30 at day 12 and later) but fail to meet the criterion in the solvent control. Furthermore these data do not indicate whether the within-treatment variability in egg production is comparable with the results from DTU.

Discussion on mortality in mass culture and sex ratio

Response to points made in discussion of test results (refer to pp. 36-38).

The authors state both here, and in the overall evaluation (*refer to page 38*), that more experience is needed before a criterion can be specified for this parameter (mortality). As discussed above, such high levels of mortality are not acceptable for a proposed regulatory test guideline and these results suggest that more experimental work is needed to investigate (and resolve) why egg hatch and larval mortality are so high.

The authors recommend removing sex ratio for the current Guideline until better understood. For the *Acartia* Guideline, the bias in sex ratio does not appear to be a critical factor as the life-cycle test starts with a relatively large number of animals per treatment (e.g. 500-600 eggs of which approximately 80% are expected to successfully hatch). Therefore, provided that you can accurately sex individual females for initiation of the egg production phase (i.e. 10 individual females per treatment), the high numbers of animals should ensure that sufficient females are

available for the egg production phase. As the authors state, many factors influence sex ratio in copepods. Biased sex ratios alone may not unduly influence the successful outcome of a test, unless they are an indication of suboptimal conditions in the cultures, which subsequently may adversely affect the outcome of the test. The results from the pre-validation work with harpacticoid copepods may shed further light on this subject.

Egg production

Refer to pages 24-27.

The discussion of egg production data (page 24) by DTU identifies problems with low numbers of females per group and high variability in egg production between individuals within a treatment. These results emphasize the need for further exploratory work to investigate the measurement of egg production in this species. For example, the authors have indicated that variation in temperature may have contributed to low egg production by control females in the fipronil experiment. The measurement of reproductive output in the life-cycle test is important, however, these results suggest that further experimental work is required to refine/improve the parameter of egg production. Such work is required before ring testing can proceed otherwise non-expert laboratories may find it difficult to achieve valid results for egg production.

Discussion on egg production

Response to points made in discussion of test results (refer to pp. 36-38).

As discussed earlier, selecting females at the time when they can be sexed is a simple solution to the problems experienced by DTU in obtaining sufficient numbers of females per treatment for the egg production phase. It was also suggested that more accurate temperature control ($\pm 1^\circ\text{C}$) is required for the measurement of egg production.

It is interesting that egg hatch and subsequent larval development of the F1 generation are followed in the mass culture and not in the isolated females. Compared with the mass culture (i.e. large numbers of eggs daily), fewer eggs are produced daily by the individual females. However, provided that it was practicable, measuring LDR in eggs from the isolated females may be preferable, in so much that you would be able to identify effects on egg production, egg hatch and subsequent larval development in the same individuals or group of individuals (n=10). You could probably obtain sufficient eggs for the LDR test by pooling the eggs produced within each treatment (n=10). Does this approach have any advantages for simplifying the experimental design?

--- 000 ---

(Tim Williams):

These comments were sent to Ole on 13 July 2005. Ole responded on 14 July 2005 and my reply is enclosed below:

Hi Ole,

Feedback on report

Many thanks for the feedback. As expected, you have already addressed most of the points raised in my review of the report. I appreciate your concern over my suggestion for more validation work and I am happy to 'qualify' this statement (here and to the wider circulation).

The data presented in the report suggested to me that further experimental work was needed to investigate some of the key technical issues. However, your subsequent response to my comments, supported by additional data not presented in the validation report, has addressed most of these technical points and explained them more fully. Your suggestion to include these comments in the final report is a good one and will inform the wider audience of the technical discussion.

Having responded to the technical comments, you (the expert laboratories) are best placed to decide whether further validation work is needed before ring testing. I look forward to reading the final report and hope that we have the opportunity to meet soon to discuss the work at 'first hand'.

Best regards,

Tim

Contact details

Dr Tim Williams
Senior Scientist
AstraZeneca UK Limited
Global SHE Operations
Brixham Environmental Laboratory
Freshwater Quarry
Brixham Devon TQ5 8BA
United Kingdom
tel: +44 1803 882882 (switchboard)
tel: +44 1803 884271 (direct line)
fax: +44 1803 882974
e-mail: tim.williams@astrazeneca.com

Response (14 July 2005) by K. Ole Kusk, DTU, to comments by Tim Williams:

Hi Tim

Thanks again for your comments. I'm a bit concerned about one of your recommendations that more validation work is needed, because I fear that if that will be the decision the whole work will be delayed

significantly or maybe even stopped. Besides I do not agree with you and with the following I hope to convince you to moderate your recommendations.

When we started on this we didn't have much experience in life-cycle tests with *Acartia*. We had concentrated on the early life stage test (LDR of F0) and had also experience with egg production of adult females.

When we designed the life-cycle test we tried to include some good test parameters know from among others the *Daphnia* reproduction test. Also the fact that we started out with one common draft guideline had some influence on the choice of parameters. But we had no experience of routine testing with the onset of egg production and we had no data on survival over a longer period. This was of course a problem when we should put up some validity criteria.

We have much more experience now because we have used the guideline not only for this pre-validation work but also in our EU project work within Compendo.

One main problem with the pre-validation work is that DHI stucked too much to their "old" procedure and performed their tests in 32 o/oo saltwater where the development runs slower - nevertheless they stopped their tests after exactly 5 days (120 hours). Thus it's not a temperature control problem but simply the fact that they should have waited longer to reach a copepodite fraction of 50 %.

DHI also performed their egg production test differently and their results cannot be compared to our results. On the other hand it gave the inspiration for a change in this part of the guideline, and we plan to design this part more like the egg production test we did before on mature females.

With respect to optimal culture conditions: I think this is really problematic. Suppose the conditions are optimal - how can we then prove this. If they are not we could work on an improvement, but this could take a very long time - and I actually think that we wouldn't be able to improve this much. Our cultures have been kept under the present conditions for years and performed well. Now and then there may be a problem - but it always turn out to be some mistakes with the preparation of media or poor food quality due to bad planning or mistakes with the algal media. We have used the synthetic media now for about 8 years without any problems. And I and Estella at DHI each have almost 20 years of experience with *Acartia*.

Larval development ratio: The problem is as mentioned that DHI used 32 o/oo media and stopped exactly after 120 hours. This is simply solved by using 20 o/oo instead and by following the given procedure for checking the development regularly before stopping the test.

Egg production failure: I have attached (see attachment) results from earlier (controls) egg production tests, where you can see that there are no problems of reaching 30 eggs per female and day. The reason here is maybe temperature but after having given it a second (and third) thought I think the low number is caused by the fact that the animals haven't reached their maximal production at day 12. With the new procedure we are going to propose we will start with 14 days old

females - and then I don't expect any problems fulfilling the criteria of 30 eggs per day and per female.

Survival of larvae: I have gone through our results from 15 recent tests and found that we need to moderate the criteria a bit. In 14 out of the 15 test we had a control survival of at least 70 %, but only 7 fulfilled the present criteria of 80 % survival. In one test out of the 15 survival was below 70 %. (See attachment - second page). Thus, a validity criteria saying less than 30 % mortality in the control would be more appropriate - and when we start with at least 50 eggs (taking also the hatchability into consideration) we still have at least $50 * 0.8 * 0.7 = 28$ animals per vessel. (By the way - when we calculate survival we subtract the number of unhatched eggs from the start number of eggs)

The mortality in the basic exposure: I have no other explanation for the high mortality than this is a short lived organism. The control mortality was confirmed by the paper by Barata et al. (2002) referred to in the report. I will try to look for more support for this but I'm not too optimistic that I can find anything. Maybe Estelle have some data? I will ask her. If you have something on this please inform me. The high mortality in the basic exposure makes this parameter unacceptable as a test parameter. On the other hand there have to be some validity criteria here to ensure that enough animals survive to have animals for the egg production test and for measuring of length. One possibility could be to put up two criteria - one for survival and one for sex ratio. Another more labour-saving criteria could be to have just one criteria saying at least 50 control females and 50 control males at the end of exposure (about day 17).

Sex ratio: The same as above - it's too uncertain to be used as a test parameter and we can actually avoid having a criteria on this by using the criteria mentioned above (50 males and females).

50 % copepodite fraction - instead of 80 %: The reason is that at 50 % copepodite fraction you have the biggest distance between an inhibited culture (or stimulated) than at any other points. I will try to find a figure which can illustrate this and include it in the final report.

I think this responds to most of your main comments - the final report will reflect at least part of this. Maybe I will include comments as Annexes in the report - though it may be difficult with those from Henrik Tyle.

I haven't finished the discussion with Estelle yet about the next version of the draft guideline, but I have included the present working paper on the subject, where you can see what we have in mind.

Best regards
Ole

Annex 1 - Table 1: Survival of hatched nauplii and copepodites in LDR tests performed at DTU from 2002-2005

Survival in LDR tests

Chemical	Generation	Survival in %
Chm1	1. gen	70
Chm2	1. gen	77
Chm2	2. gen	81
Chm3	1. gen	94
Chm3	2. gen	86
Chm4	1. gen	99
Chm4	2. gen	66
Chm5	1. gen	71
Chm5	2. gen	86
Chm6	1. gen	87
Chm6	2. gen	79
Chm7	1. gen	71
Chm7	2. gen	79
Chm8	1. gen	73
Chm8	2. gen	86

Table 2: Control egg production in egg production tests performed at DTU in the period from 1998-2001. Groups of females were exposed and the numbers of eggs produced in the 3rd and 5th day after onset of exposure were counted.

Egg production in control experiments 1998 - 2001

Year	Month of year	No. of replicates	Average egg production per female per day	Testcompound
2000	1	8	59.1	TetraBr Bis A
1998	1	10	55.7	NPE
2000	2	6	72	DES
1998	2	9	48.5	Bis A
2000	2	8	44	MT
1998	3	7	46.9	Estradiol
2000	3	8	42.4	Flutamide
2000	3	8	59	EthinylEstradiol
1999	4	8	47	Lundtofte STP
1998	4	8	63.2	Testosterone
1999	5	8	46.4	Avedøre sp.v.
2003	7	8	42.8	Fenarimol
1999	8	8	65.3	Ecdysone
1999	8	8	58.9	Progesterone
1999	9	8	50.1	Estrone
1999	9	6	51.1	JH3
1999	10	8	53.6	EthinylEstradiol
1999	11	8	66.2	K2Cr2O7
1999	11	8	49.7	DCA
1998	11	6	52	3,5-DCP
			Average	St.dev
			53.7	8.44