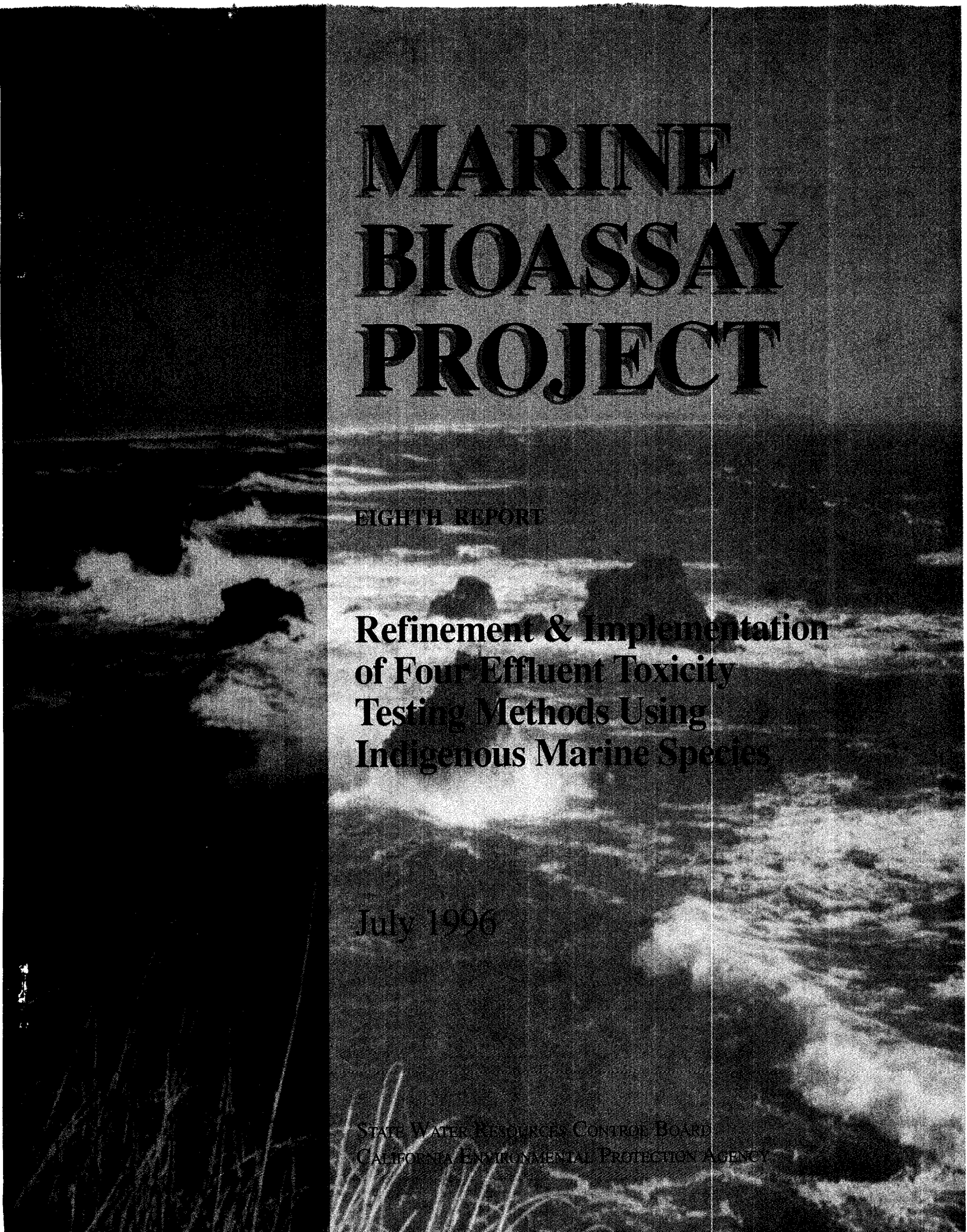


MARINE BIOASSAY PROJECT

EIGHTH REPORT

Refinement & Implementation of Four Effluent Toxicity Testing Methods Using Indigenous Marine Species

July 1996

STATE WATER RESOURCES CONTROL BOARD
CALIFORNIA ENVIRONMENTAL PROTECTION AGENCY

STATE OF CALIFORNIA
Pete Wilson, Governor

CALIFORNIA ENVIRONMENTAL PROTECTION AGENCY
James M. Strock, Secretary

STATE WATER RESOURCES CONTROL BOARD
P.O. Box 100
Sacramento, CA 95812-0100
(916) 657-1247

John Caffrey, Chairman
John Brown, Vice Chair
Marc Del Piero, Member
James Stubchaer, Member
Mary Jane Forster, Member

Walt Pettit, Executive Director
Dale Claypoole, Deputy Director

MARINE BIOASSAY PROJECT

EIGHTH REPORT

July 1996

John W. Hunt, Brian S. Anderson, Shirley Tudor
Institute of Marine Sciences, University of California,
Santa Cruz, CA 95604

Mark D. Stephenson and H. Max Puckett
California Department of Fish and Game
Marine Pollution Studies Laboratory, 34500 Coast Route 1,
Monterey, CA 93940

Francis H. Palmer and Matthew W. Reeve
State Water Resources Control Board, P.O. Box 100, Sacramento, CA 95801

STATE WATER RESOURCES CONTROL BOARD
CALIFORNIA ENVIRONMENTAL PROTECTION AGENCY

Acknowledgments

The authors thank the staff of the Marine Pollution Studies Laboratory for their assistance with all phases of this project: Steve Clark, Matt Englund, Michelle Hester, Steve Osborn, Bryn Phillips, Witold Piekarski, Kelita Smith, and Lisa Weetman. For their guidance and many helpful suggestions, we are grateful to the members of the Marine Bioassay Project Scientific Review Committee: Steven Bay, Southern California Coastal Water Research Project; Tom Dean, Coastal Resources Associates; Jo Ellen Hose, Occidental College; and Phil Oshida, U.S. EPA. Several laboratories were involved in the interlaboratory comparisons; the cooperation and suggestions of all of the participants contributed to the development of the toxicity tests. Jon Goetzl, Gary Ichikawa, Jim Kenihan, and Kim Paulson of the California State Mussel Watch Program performed the chemical analyses of reference toxicant test solutions; their cooperation is appreciated.

This report was prepared through agreement with the California State Water Resources Control Board (Contract No. 2-041-250-1) and subcontract (FG-0534) with the California Department of Fish and Game (Project Managers: Mark D. Stephenson and H. Max Puckett) to the University of California, Santa Cruz (Principal Investigator: Dr. Ron Tjeerdema). This does not signify that the contents necessarily reflect the views and policies of the California State Water Resources Control Board, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

TABLE OF CONTENTS

List of Figures

List of Tables

Executive Summary

The goal of the Marine Bioassay Project (MBP), authorized by the State Water Resources Control Board (State Board) in 1984, is to protect California's ocean resources by determining the impacts of toxic waste discharges on marine waters. The Project's primary objectives are (1) developing and validating of critical life stage tests to measure the toxicity of these discharges, and (2) providing technical support to State and Regional Board staff and to outside laboratories performing the toxicity test methods listed in the California Ocean Plan.

In its 13 year history, the MBP has developed four critical life stage protocols using species indigenous to California's coastal waters. These reflect a newer generation of test methods designed to estimate more subtle long term adverse effects of waste discharges that may ultimately damage populations of important marine species. The four toxicity test protocols use an alga (giant kelp), a fish (topsmelt), and two invertebrates (red abalone and Pacific mysid crustacean). The giant kelp (*Macrocystis pyrifera*) and red abalone (*Haliotis rufescens*) protocols are currently on the 1990 Ocean Plan approved list of critical life stage protocols for use in NPDES compliance monitoring. The mysid (*Holmesimysis costata*) and topsmelt (*Atherinops affinis*) protocols were added to the list with the adoption of the 1996-1997 Ocean Plan amendments by the State Board. The 1997 Plan will be effective after approval by the Office of Administrative Law. All of the MBP toxicity test methods are published in

the Procedures Manual for Conducting Toxicity Tests Developed by the Marine Bioassay Project (96-1WQ).

Two State agencies and one university work cooperatively in operation of the MBP. The Department of Fish and Game's (DFG) Marine Pollution Studies Laboratory (MPSL), serves as the research facility. The University of California at Santa Cruz provides the principal investigators and staff to conduct the research. The State Board oversees and provides funding for the Project by contracting with the DFG to develop and conduct tasks specified in the contract's agreement. DFG in turn subcontracts with U.C. Santa Cruz to provide staff and operate the Project.

The MBP staff has worked closely with the EPA in adapting MBP protocols for inclusion into the new EPA West Coast Methods Manual (EPA/600/R-95/136). This EPA manual contains seven critical life stage protocols including the four developed by the MBP.

Regulatory Background

Development of toxicity test protocols to estimate long-term effects of waste discharges is consistent with both federal and state requirements. In 1984, the United States Environmental Protection Agency (EPA) issued a national "Policy for the Development of Water Quality-Based Limitations for Toxic Pollutants" (49 CFR, No. 48, March 9, 1984). This policy outlined a technical approach for controlling discharge of toxic substances through the federal

system of discharge permits. In addition to meeting numerical standards for individual chemicals, the policy requires EPA and the States to use biological testing to complement chemical testing. Biological testing also provides information not available for chemical testing. For example, it incorporates bioavailability and interactions in complicated mixtures of toxic materials.

In 1986, AB 3500 added Section 13170.2 to the California Water Code. In addition to mandating triennial review of the California Ocean Plan, Section 13170.2 requires the State Board to develop and adopt toxicity test protocols. Ocean discharges of 100 million gallons per day or more have been required in their permits to use their toxicity test protocols for monitoring complex effluents since January 1, 1991. This requirement was extended to smaller dischargers on January 1, 1992. Section 3 of AB 3500 expressed legislative intent that the organisms used in testing be representative marine species.

> "If the State Water Resources Control Board determines through its Marine Bioassay Project that a multispecies toxicity testing program with representative marine species for monitoring complex effluent discharges is appropriate, the state board shall use the multispecies toxicity testing program with representative marine species in adopting the toxicity test protocols specified in Section 13170.2 of the Water Code."

On March 19, 1990 the State Board adopted a series of amendments to the California Ocean Plan. These amendments included the addition of a chronic toxicity objective for protection of

marine aquatic life. The State Board also adopted a list of seven toxicity test protocols deemed sufficiently developed for measuring compliance with the chronic toxicity objective. Included on this list of seven are two protocols developed by the MBP, the giant kelp and the red abalone 48-hour toxicity tests. The MBP has since developed two additional critical life stage tests using the Pacific mysid (*Holmesimysis costata*) and the topsmelt (*Atherinops affinis*). These toxicity protocols were added to the Ocean Plan list with the adoption of the 1996-1997 Ocean Plan amendments by the State Board.

In practice, toxicity requirements in a discharge permit are expressed in toxic units (TU). A TUc is defined as 100 divided by the No Observed Effect Concentration (NOEC):

$$TUc = 100/NOEC$$

The NOEC is defined as the maximum percent concentration of effluent, or any water being tested, that does not result in any observed effect on test organisms. Permits would usually require that no sublethal toxicity be observed at a concentration lower than those present within an outfall's designated mixing zone (the "zone of initial dilution"). For example, if a discharger has an outfall design that provides 99:1 dilution, then no toxicity should be observed in effluent diluted to one percent. The discharge permit would require that the effluent toxicity limit be 100 toxic units or less.

$$TUc = 100/1 = 100$$

Project History and Accomplishments

The Marine Bioassay Project is designed as a multiple phase program to develop and implement short-term tests for toxicity measurement of complex effluents discharged to the ocean. Actual laboratory work is conducted at the California Department of Fish and Game's (DFG) Marine Pollution Studies Laboratory located south of Monterey. To date, seven phases of the Marine Bioassay Project have been completed; chapters of this report describe work performed during the seventh phase. This seventh phase is two-pronged: (1) refinement and (2) implementation of the four MBP protocols. Improvement of test methods has included (1) evaluating the sensitivity, precision, and logistical feasibility of the *Holmesimysis costata* protocol, (2) examining the use of small volume containers for toxicity tests involving mollusc larvae, (3) evaluating the influence of salinity on copper toxicity to larval topsmelt (*Atherinops affinis),* and (4) determining the biological significance of the short term (48) hour giant kelp (*Macrocystis pyrifera*) test. Implementation has entailed training of technical staff and providing technical support for participating laboratories.

Phase One (November 1984-February 1986)

During the first phase, efforts were made to obtain wide-spread participation in developing the scope of the project. Initially, a draft report was prepared that described a number of potential marine toxicity test species, recommended twelve of these as most suitable, and presented appropriate protocols for each of the twelve. In March 1985, the draft was sent for

review to a number of potentially interested agencies (NOAA and DFG), ocean dischargers in southern California, the Southern California Coastal Water Research Project, and a number of individual scientists.

A workshop to discuss the draft report and outline the project's scope was held on April 29, 1985 at the offices of a major ocean discharger, the County Sanitation Districts of Orange County. The purpose was to discuss the proposed toxicity test species and protocols and address questions raised by reviewers of the draft report. Over 50 people attended the workshop and general session and then participated in one of the five sub-committee meetings. The MBP's First Report included a summary of the workshop proceedings in addition to the species descriptions from the pre-workshop draft.

A separate outcome of the workshop was the establishment of a Scientific Review Committee, composed of a small group of outside experts to discuss progress and provide guidance for the Marine Bioassay Project. The first meeting was held in June 1985, and meetings have continued approximately twice a year. The Committee has recommended a number of significant mid-course corrections that have been implemented by the MBP staff. Overall, a major accomplishment in these recommendations has been to refocus daily work on the primary objective: development of short-term protocols for use in performing toxicity tests on complex effluent discharged to the ocean.

Four important laboratory tasks were also completed during Phase One: (1) the Department of Fish and Game laboratory was extensively refurbished for animal culture and rearing of marine species, (2) methods were developed for maintaining and spawning selected marine species, (3) a mobile laboratory was purchased and used to conduct aquatic toxicity tests, and (4) range-finding and definitive tests were developed on two toxicants (pentachlorophenol and endosulfan) of immediate concern to the State Board. This work is described in the MBP's Second Report, (May 1986).

Phase Two (March 1986-October 1987):

In Phase Two, three new short-term protocols were developed after repeated testing using zinc as a reference toxicant. In addition, longer term reference toxicant tests were used with each species to calibrate the relative sensitivity of the short-term test protocols. All three short-term test protocols developed were static tests; that is, the test solutions are not changed during testing. Each protocol measured a different effect or endpoint.

These protocols, designed to estimate chronic toxicity of discharge to ocean waters, utilized sensitive life stages of three marine species: the red abalone *Haliotis rufescens*; a mysid shrimp, *Holmesimysis costata*; and the giant kelp, *Macrocystis pyrifera*. After some refinement, preliminary testing with the three protocols was performed on complex effluents from two representative municipal treatment plants.

The short-term larval abalone toxicity test protocol is a 48-hour test in which abnormal shell development is the endpoint used as the measured effect of toxicity. The short-term giant kelp toxicity test is a 48-hour test that measures two different endpoints: zoospore germination and growth of the germination tube. The short- term mysid toxicity test is a 7 day test with endpoints of survival and growth to juvenile mysids.

Phase Three (November 1987-December 1988)

During this phase, the abalone, kelp, and mysid shrimp tests developed during Phase Two and described above were further refined using complex effluent from two large municipal ocean dischargers. In addition, preliminary tests were conducted using a fish species, topsmelt *Atherinops affinis*. The project's fourth report provides detailed descriptions of work completed in Phase Three.

Phase Four (January 1989-December 1989)

During phase four, the giant kelp, red abalone, and topsmelt toxicity tests developed in previous phases were refined by testing with complex effluents. In addition to performing toxicity tests of complex effluents, a manual entitled "Procedures Manual For Conducting Toxicity Tests Developed By The Marine Bioassay Project" was prepared for the red abalone, giant kelp, mysid shrimp, and topsmelt protocols. This manual is available from the State Board (Report 90-10WQ).

Phase Five (January 1990-December 1990)

In Phase Five, complex effluent tests were conducted and toxicity protocols developed using topsmelt, a fish native to the California coast. Also, a sub-lethal endpoint was established for the mysid shrimp protocol, and additional research was conducted in the areas of complex effluent testing and interlaboratory verification. Finally, technical support was provided by the MBP staff to laboratories using MBP protocols. This support included training of laboratory technicians, providing broodstock organisms, detailing specific requirements for test acceptability of individual protocols and performing interlaboratory tests with discharges exceeding 100 million gallons per day of complex effluent.

Phase Six (January 1991-June 1993)

Phase Six completed protocol development by the project and continued implementation of the marine toxicity testing program.

Major tasks completed include:

1. Completing the test protocols for the mysid shrimp and the topsmelt.
2. Providing additional technical training and support for dischargers and consulting laboratories.
3. Evaluating the biological significance of the short term (48 hr) red abalone routine test, by comparing it to both long term and exposure-recovery tests.

4. Investigating the use of marine organisms for measuring ambient toxicity from non-point sources.

5. Evaluating the applicability of sodium azide as a replacement toxicant for copper when conducting toxicity studies with the giant kelp *Macrocystis pyrifera*.

Phase Seven (July 1993- March 1996)

Phase seven (eighth report) is organized into six sections. Section 1 focuses on topsmelt, *Atherinops affinis*, one of the most abundant fish species in central and southern California estuaries. Section 2 describes work with the mysid shrimp, *Holmesimysis costata*, a crustacean which inhabits the upper canopy of giant kelp forests. Section 3 describes work with the red abalone, *Haliotis rufescens*, a large gastropod mollusc indigenous to California and distributed throughout the State coastal waters. Section 4 focuses on the giant kelp, *Macrocystis pyrifera*. California kelp forests harbor a rich diversity of marine life and are an important source of primary production to the nearshore marine ecosystem. Section 5 describes the use of Minimum Significant Difference (MSD) criteria to address differences in test sensitivity among MBP protocols. Section 6 summarizes the several remaining tasks the MBP has completed. These tasks include (1) promoting use of the various toxicity test protocols by establishing a network of reliable test organism suppliers (supply has been established for three of the four MBP protocols), (2) facilitating development of a fully operable toxicity database/statistics package for use statewide, (3) conducting a workshop as part of an ongoing statewide training for Whole Effluent Toxicity (WET) testing for EPA,

State and Regional Board staff, (4) participating in an EPA-sponsored workshop on bioremediation metabolite toxicity, and (5) participating in several meetings and workshops on statistical issues concerning toxicity test data.

Topsmelt studies

The topsmelt studies examined salinity effects on larval sensitivity. This was achieved by examining the influence of salinity on copper uptake and the effect of copper toxicity on larval osmoregulation. Osmoregulation is the process by which an animal can control water levels and salt concentrations within its body. Marine organisms such as topsmelt are able to maintain a higher freshwater to salts ratio internally than the surrounding marine environment. Differences in salinity did not appear to significantly influence copper uptake in larval topsmelt. However, the copper data did indicate an increase in larval mortality with decreased salinity. This result was probably due to the combination of copper toxicity and the physiological stress induced by lower salinity. Study results demonstrated that larval topsmelt are amenable to toxicity testing at estuarine salinities varying from 10 to 33 parts per thousand.

Pacific mysid studies

The mysid experiments evaluated the sensitivity, precision, and logistical feasibility of the 7-day growth and survival toxicity test using *Holmesimysis costata*, an indigenous west coast

marine crustacean. Studies were conducted to estimate inter- and intralaboratory precision using Coefficient of Variation (CV) values, and to evaluate weight versus length as measurements of growth. The CV is a statistic often used to measure the precision of different toxicity tests. It is the standard deviation expressed as a percentage of the mean. In this instance, the CV is used to measure the precision among laboratories conducting the *Holmesimysis costata* toxicity test.

The intra- and interlaboratory CV values for the *H. costata* toxicity test were comparable to CV values for intra- and interlaboratory precision estimates for analytical chemistry techniques used in permit monitoring. Six tests using the reference toxicant zinc sulfate conducted at different times in the same laboratory (intralaboratory) produced a CV of 25% among test's using Lethal Concentration 50s (LC50) as an endpoint. LC50 is the concentration of a toxicant in water which is lethal to 50% of the test organims. The interlaboratory precision for *H. costata* data was determined to have CV values ranging from 7 to 30%, with a mean of 14%.

Two growth endpoints, change in weight and change in length of the test organisms, were compared in five effluent tests. Both Inhibition Concentration 10 (IC10) and IC25 values were lower for the weight endpoint than for length in four of the five tests, indicating greater sensitivity for the weight measurement. IC10 is defined as the concentration of a toxicant in water which causes some adverse effect on 10 % of the test organisms. However, variability among replicates was lower in the measurements of length data, resulting in lower Minimum

Significant Difference (MSD) values. The MSD is a statistic commonly used to measure the power of toxicity tests and by establishing criteria, it can also be used as a method to control toxicity test power. Toxicity test power is a statistic used to detect a certain level of effect. A toxicity test with good power would have a low probability of giving erroneous results (a low probability of stating that a solution does not contain toxicity when there really is toxicity present). The lower MSD values in the mysid test length endpoint allow smaller differences from control values to be statistically significant, even when point estimates such as an IC10 indicate a lower overall effect level. Given that test endpoints should measure biologically significant effects, the weight measurement appears more likely to result in NOEC values indicative of reduced growth as opposed to reduced variability. Overall, these data obtained during the development of the mysid protocol support using the weight measurement as the preferred means of determining the growth endpoint for the test.

Red abalone experiments

The red abalone experiments investigated the use of smaller test containers for toxicity testing. This technical development has a number of advantages over conducting tests in large beakers, advantages which include greatly reduced sample volume requirements, logistical benefits, and, most importantly, the elimination of subsampling error and bias in the analysis of toxic effects.

In two experiments, there were significant differences in embryo-larval development among tests conducted in different types of containers. The differences are indicated by Effective Concentration 50 (EC50) values for the respective dose-response curves. EC50 is the concentration of a toxicant in water which causes a response in fifty percent of the test organisms. In the two experiments evaluating effects of container size, tests conducted in small vials resulted in lower EC50 values than tests in beakers, indicating greater toxicity observed in small containers. Previous comparisons of data from numerous non-concurrent reference toxicant tests indicated no significant difference in toxicity among test container types. Further testing would be necessary to investigate the significance of the differences observed in the presently described experiments.

Giant kelp studies

The giant kelp exposure recovery experiments investigated biological significance of endpoints measured in the 48 hour (short-term) *Macrocystis* germination and growth protocol. These experiments were designed to determine whether vegetative effects (percent germination and kelp germ tube length) observed in short-term exposures are indicative of longer-term reproductive effects, and whether cultures are capable of recovery when transferred to toxicant free media. The effects of duration of initial exposure on eventual recovery were also examined. The toxicants used in the study were copper and sodium azide. Although preliminary, the study results suggest that the initial exposure duration determines whether kelp gametophytes exposed to sodium azide are competent to recover and

successfully reproduce when transferred to uncontaminated media. Cultures exposed to azide and copper for 48 hours were generally able to recover, while those exposed to azide for 96 hours were not. The initial exposure concentration apparently affected the ability of the cultures to recover. Studies with additional toxicants are necessary to determine whether these results are consistent for compounds having different modes of action.

Topsmelt supply

MBP staff was successful in establishing a reliable supply of topsmelt, *Atherinops affinis*, for toxicity testing. Two of the nation's largest test organism suppliers are now culturing topsmelt, the suppliers are Aquatic Indicators in St. Augustine, Florida, and Aquatic Biosystems in Fort Collins, Colorado. MBP staff also participated in an EPA-sponsored workshop on chronic toxicity at the EPA's Region 9 Richmond Field Station in June, 1994, as part of the ongoing statewide training for WET testing. In attendance were State and Regional Board staff.

Toxicity database

The MBP staff continued to work with a private contractor on the development of TOXIS, a toxicity database/statistics package designed for entering, analyzing, and retrieving WET test data. MBP staff used the program on a test basis to help facilitate development of a fully operable program for statewide use. Project staff have also participated in workshops such as

the 1994 National Society of Environmental Toxicology and Chemistry meeting, to address statistical issues regarding toxicity test data analysis.

Section 1

Topsmelt (*Atherinops affinis*) Experiments

Influence of salinity on copper toxicity to larval topsmelt *Atherinops affinis*

INTRODUCTION

A 7-day growth and survival toxicity test protocol for the topsmelt, *Atherinops affinis*, has been developed by the California State Water Resources Control Board as part of an effort to develop protocols to estimate chronic toxicity to species indigenous to the west coast of the United States (California State Water Resources Control Board 1996; U.S. EPA 1996). In previous work with topsmelt we have demonstrated that this species is amenable to toxicity testing at estuarine salinities. However, results of toxicity tests conducted at salinities from 5 to 33‰ showed that toxicity of both copper chloride and sodium azide increased at lower salinities (Anderson et al. 1994).

In an effort to explain salinity effects on topsmelt larval sensitivity, additional experiments investigated the influence of salinity on copper uptake and the effect of copper toxicity on larval osmoregulation. The influence of salinity on copper chloride (cupric chloride; $CuCl_2$) toxicity was investigated because copper is a commonly used reference toxicant and ubiquitous contaminant (Weber *et al.* 1988, Anderson *et al.* 1991and 1994), and because cupric ion complexation and bioavailability are affected by organic ligand concentrations which may vary with salinity. There have been few studies on the influence of salinity on copper toxicity to larval fish. Results of this study are discussed in terms of the interaction of osmotic stress and copper toxicity, and in the context of using protocols for topsmelt and other euryhaline fishes to assess ambient toxicity in receiving waters of varying salinity. The research was included in a recent Marine Bioassay Project publication (Anderson et al. 1995).

METHODS

Facilities

All experiments were conducted in February, 1994 at the California Department of Fish and Game's Marine Pollution Studies Laboratory (MPSL) located at Granite Canyon in Monterey County, CA. Background concentrations of total copper in the laboratory's filtered natural seawater were determined to be below the detection limit of 3 µg/L in this study.

The influence of salinity on copper uptake and larval osmoregulatory function were determined during copper exposures using larvae which had been acclimated to different salinities through embryonic and early larval development.

Effects of Copper on Osmoregulation

To investigate the effect of copper on larval osmoregulation, early blastula-stage embryos were transferred from the culture tank (34‰) directly to culture tubes at salinities of 10, 17, 25 and 34‰ until hatching. After 10 days, hatched larvae were then exposed to copper at the respective salinities for 7-d. Copper concentrations for these exposures were comparable to the copper LC50 at each salinity (40, 70, 130, and 200 µg/L in 10, 17, 25, and 34 ‰, respectively). After 7-d, tissue and culture media osmolality were measured using a Wescor Vapor Pressure Osmometer (model 5500). The larvae had insufficient blood volumes for direct measurement of osmolality, so approximately 100 larvae were removed, rinsed with deionized water, thoroughly blotted dry, and placed into a test tube for homogenization with a clean glass rod. A 24-mg subsample of the resulting homogenate was then placed in the osmometer for measurement; three separate measurements were taken for all fish at all salinities. Two standards bracketing expected sample and media values were analyzed with each set of

three tissue samples. Tissue osmolalities from exposed and control fish were compared using t-tests.

Effects of Salinity on Copper Uptake

A separate experiment investigated whether copper uptake varied with salinity. As described above, early blastula-stage embryos were transferred from the culture tank (34‰) directly to culture tubes at 10 to 34‰ until hatching. After 9 days in culture, the hatched larvae were then exposed to 32 µg/L of copper for 7 days in media salinities of 10, 17, 25, and 34 ‰ (n = 3 replicate exposures/treatment). Control fish were cultured at 34‰ in the absence of copper. After 7 days the larvae were rinsed with distilled water, blotted dry and frozen. Fish tissue samples were heated on a hot plate for five hours and digested with concentrated 4:1 nitric:perchloric acid in a Teflon® vessel. The liquid digestate was then diluted with 18 MΩ deionized water to a final volume of 20 ml, and measured using a Perkin-Elmer Model 3030 Zeeman atomic absorption spectrophotometer (detection limit = 0.04 µg/gram) using methods described in Lauenstein and Cantillo (1993). Copper concentrations were expressed as total copper (Cu^{2+}) on a wet weight basis. Tissue copper concentrations from fish exposed at different salinities were compared using ANOVA.

Test Media Chemical and Physical Measurements

Physical/chemical parameters of dissolved oxygen, pH, salinity and temperature were monitored in each test concentration at the beginning and end of each experiment, and before one renewal. Dissolved oxygen concentrations and pH were measured using an Orion Model SA 520 pH/millivolt meter. Salinity was measured using a Reichert-Jung model 10419 temperature-compensated refractometer. Temperature was monitored daily with a mercury thermometer. Water quality parameters were within acceptable limits in all experiments. The mean temperature was 21 ± 1°C, dissolved oxygen was

always greater than 80% saturation, and measured salinity was within ± 1.5‰ of nominal values. The pH values were always near 8.0 (the range across all experiments was 7.72-8.54).

To verify copper concentrations in the exposure media, one random sample of each test concentration was taken at the beginning and end of each test and analyzed using atomic absorption spectrophotometry. Analytically measured copper concentrations varied 9.8% ± 4.4% (mean variation ± standard deviation) from nominal concentrations, which is within the accuracy of the technique. All copper values are reported as nominal values.

RESULTS

Effects of Copper on Osmoregulation

Larval tissue osmolality was greater in non-exposed fish than in copper-exposed fish at all salinities (Table 1). This difference was statistically significant at all but the 10‰ salinity (Table 1). The relative increase in tissue osmolality in the non-exposed larvae was 5%, 21%, 20%, and 16% at 10, 17, 25, and 34‰ salinities, respectively. The isosmotic point (where the control culture media osmolality equals the control tissue osmolality; Hoar 1975) was determined graphically to be approximately 12‰ for larval topsmelt (400 mmol/kg; data not shown). Note that the tissue osmolality of copper-exposed larvae approached the osmolality of the non-exposed larvae at 10‰ (Table 1).

Copper Uptake

Differences in salinity did not appear to significantly influence copper uptake. After a 7-day exposure to 32 µg/L copper, tissue concentrations increased 10-fold from a mean of 0.4 mg/L in the control larvae to approximately 3.4 mg/L in the fish exposed at

the 4 different salinities. Tissue copper concentrations were 3.8, 3.1, 3.3, and 3.2 µg/L at 10, 17, 25, and 34 ‰, respectively, and were not significantly different (ANOVA, $p = 0.065$).

Table 1. Tissue osmolality (mmol/kg) in topsmelt larvae exposed to copper chloride at four salinities. Larvae were exposed to concentrations approximating the copper LC50 at each salinity. Each tissue osmolality value is the mean (±s.d.) of three replicates of larval tissue samples at each of the four salinities.

Salinity	**10‰**	**17‰**	**25‰**	**34‰**
Copper Exposure	**40 µg/L**	**70 µg/L**	**130 µg/L**	**200 µg/L**
Copper exposed larvae	371.33 (13.87)	379.33 (7.09)	383.33 (8.50)	429.67 (15.14)
Control larvae	389.00 (15.52)	481.67 (17.01)	480.33 (9.02)	500.33 (14.05)
t-test p value (α=0.05)	p = 0.2155	p = 0.0007	p = 0.0002	p = 0.0041

DISCUSSION

This work demonstrates that larval topsmelt are amenable to toxicity testing at estuarine salinities from 10 to 33‰. Control survival at all salinities in all experiments was 100%. This salinity tolerance, year-round culturability (Anderson et al. 1994), and relative sensitivity to a variety of toxicants (Hemmer et al.1992, Middaugh and Anderson 1994) suggest that the topsmelt is appropriate for toxicity assessment in euryhaline situations.

The copper data suggest that the increase in larval mortality with decreased salinity was probably due to increasing physiological challenge of osmoregulation combined with copper toxicity. The reduction in tissue osmolality of larvae exposed to copper relative to non-exposed larvae suggests copper interfered with larval osmoregulation, resulting in reduced ion concentrations (Table 1). Several researchers have found evidence that copper inhibits enzyme systems necessary for osmoregulation (Thurberg and Collier 1973, Stagg and Shuttleworth 1982). Using fungi as a model system, Corbett *et al.* (1984) suggest that copper interferes with the pyruvate dehydrogenase system integral to the production of ATP; however the interference of copper with this enzyme system has not been described in fish larvae. A reduction in ATP would likely impair ATP-dependant active transport processes necessary for maintaining osmotic balance (Heath 1987). Crespo and Karnaky (1983) found that copper partially inhibited the ATP-dependant Na^+ pump as well as chloride cell function in gill epithelia of *Fundulus heteroclitus*. Inhibition of chloride cells might also explain our observation of reduced osmolality in topsmelt larvae. Bouquegneau and Gilles (1979) reported that the primary physiological mechanism for the interaction of salinity with metal toxicity in invertebrates is disruption of osmoregulation, with divalent metal ions competing with calcium and magnesium at uptake sites. Heath (1987) suggested

that, in most cases, copper reduces blood osmolality in fresh water fish and increases osmolality in salt water fish; euryhaline topsmelt appear to respond more like fresh water fish in this regard.

It is also possible that effects of copper on osmoregulation were secondary. Cardeilhac *et al.* (1979) found that copper ions caused physical damage to the gills of sheephead (*Pimelometopon pulchrum*). Copper disrupted the membrane integrity in gill lamellae, leading to an elevation in the concentration of plasma potassium which ultimately disrupted cardiac function. In the present study, if there was physical damage to the gills, it apparently resulted in decreased ion concentrations.

Copper uptake did not vary with salinity, and is not likely to be responsible for salinity effects on toxicity. There was a ten-fold increase in tissue copper concentration in copper-exposed fish relative to the controls at all salinities (ANOVA, $p = 0.065$). The exposure concentration (32 µg/L) was below the EC_{50} for all salinities and resulted in similarly low larval mortality at all salinities (data not shown). It is possible, however, that had higher copper concentrations been used, a difference in copper uptake might have been observed. The observed similarity in copper uptake across salinities is consistent with fish physiology in hyposmotic conditions. In order to maintain osmotic balance, drinking rates in euryhaline fish generally decrease with lower external salinity (Heath 1987, Levitan and Taylor 1979). It would be expected that as drinking rate decreased, fewer copper ions would cross the intestinal wall. Although there may have been a greater concentration of free cupric ions at lower salinity due to a reduction in ligand concentrations (see below), a reduction in drinking rates would have offset this difference.

Our results are consistent with those of other studies that have investigated the interactive relationship between salinity and heavy metal toxicity in estuarine and marine organisms. Studies with fish embryos and larvae using cadmium as a representative metal have found increased toxicity at lower salinities (Alderlice *et al.* 1979, Voyer *et al.*

1979, Voyer *et al.* 1982). Studies with crustacea have also found increased sensitivity to heavy metals as salinity decreased (Jones 1975, Bjerregaard and Vislie 1986). Although the primary mechanism depends upon the metal in question, the prevailing conclusion of most studies is that metal toxicity results in disruption of osmoregulation.

It is likely that osmotic stress was coupled with increased cupric ion bioavailability at lower salinity in these experiments. Coale and Bruland (1988) found that in the upper 200 m of northeast Pacific waters, greater than 99.7% of copper (Cu^{2+}) is bound by strong organic complexes. Engel *et al.* (1981) have shown that copper is bound primarily by organic ligands in fresh- and brackish-water systems. In our experimental system, seawater was diluted to the lower salinities using distilled water, a method commonly used to adjust receiving and dilution water salinities in toxicity tests (Weber *et al.* 1988). This probably diluted organic ligand concentrations, resulting in increased copper bioavailability, and toxicity (Sunda and Guillard 1976, K. Coale pers. comm.). Variable pH was presumably not a factor in copper speciation in this experiment because the pH remained within a range of 7.7 to 8.5 at all salinities and copper concentrations. While dilution water complexation capacity was not strictly controlled in these experiments, the intent was to assess test performance using natural seawater adjusted to environmentally relevant salinities. This is the recommended practice in effluent and receiving water toxicity testing programs for which this protocol was designed. The conclusion that cupric ion concentration was inadvertently increased through dilution of organic ligands is only speculative, however, because free copper was not measured. Also, salinity and dissolved organic ligands may not necessarily be directly related to each other in nature. A second effect of diluting seawater with distilled water is the dilution of ions necessary for osmoregulation (eg. Na^{+}, Ca^{2+}, K^{+}, Cl^{-}, Mg^{2+}). With fewer ions present at lower salinities, more energy may be required to maintain proper osmotic balance. This added stress may have exacerbated the toxic effects of copper.

In a review of the effects of salinity on toxicity, Sprague (1985) suggested that euryhaline fish in general may be more effective in dealing with pollutants at their isosmotic point, possibly due to a reduction in uptake rate. Similarly, in separate studies with two different euryhaline mysid species, McClusky and Hagerman (1987) and DeLisle and Roberts (1988) found that maximum tolerance to cadmium occurred for each species at its isosmotic point. Our results indicate that this generalization is not consistent for copper and azide toxicity to topsmelt larvae. The isosmotic point for topsmelt is approximately 12‰, but our experiments indicate that both copper and azide toxicity was greater there than at higher salinities.

Although topsmelt are able to tolerate lower salinities (as indicated by 100% control survival from 10 to 34‰), this species is apparently physiologically better adapted to more marine conditions. In general, topsmelt spend most of their life at marine salinity; they spawn in harbors and estuaries during the drier summer months then migrate offshore in the fall. Although some young fish may remain, adult fish are generally not present in estuaries during the winter months when hyposmotic conditions might occur. Length and weight of larval fish in the previous study were smaller with decreasing salinity, providing further evidence of stress at lower salinities (see Anderson et al.1995, Table 3). When one stress alone (salinity) has no obvious toxic effect, but in combination with another stress (chemical) it enhances the toxicity of that chemical, the relationship is termed potentiative. Although there was 100% survival in all of the controls in the experiments reported here, reduced salinities apparently potentiated the toxicity of both copper and azide. It is possible that had we increased the range of salinity conditions tested, chemical sensitivity might have increased under hyperosmotic conditions as well.

These results demonstrate the importance of salinity in toxicity testing using estuarine species. Our results with azide and copper support previous research indicating that toxicity can be significantly affected by the interaction between salinity and chemical

stress. For industrial and municipal effluents discharged into estuarine systems, the salinity of the receiving water may significantly affect whether a given effluent is toxic to estuarine organisms. Salinity should also be considered in the design of toxicity tests used to determine whether a given discharge conforms to state and federal chronic toxicity requirements.

REFERENCES

Anderson BS, Middaugh DP, Hunt JW, Turpen SL (1991) Copper toxicity to sperm, embryos, and larvae of topsmelt (*Atherinops affinis*) with notes on induced spawning. Mar. Environ. Res. 31: 17-35.

Anderson BS, Hunt JW, McNulty HR, Stephenson MD, Palmer FH, Denton DL and Reeve M. . (1994) Marine Bioassay Project Seventh Report: Refinement of effluent toxicity testing protocols for four marine species. Water Resources Control Board Report No. 94-2WQ. Sacramento, CA. 124 pp.

Anderson BS, Hunt JW, McNulty HR, Martin M (1994) Off-season spawning and factors influencing toxicity test development with topsmelt (*Atherinops affinis).* Environ. Tox. and Chem. 13: 479-486.

Anderson BS, Hunt JW, Piekarski WJ, Phillips BM, Englund MA, Tjeerdema RS and Goetzl J.D. (1995). Influence of salinity on copper and azide toxicity to larval topsmelt *Atherinops affinis* (Ayres). *Arch. Environ. Contam. Toxicol.* 29, 366-372.

Alderdice DF, Rao TR, Rosenthal H (1979) Osmotic responses of eggs and larvae of the Pacific herring to salinity and cadmium. Helgolander wiss. Meeresunters 32: 508-538.

Bouquegneau JM, Gilles R (1979) Osmoregulation and pollution of the aquatic medium. In: R. Gillees (ed) Mechanisms of Osmoregulation. J. Wiley, New York, pp. 563-580.

Bjerregaard P, Vislie T (1986) Effect of copper on ion-and osmoregulation in the shore crab *Carcinus maenas.* Marine Biology 91:69-76.

Cardeilhac. PT, Simpson CF, Lovelock RL, Yosha SR, Calderwood SF, Gudat JC (1979). Failure of osmoregulation with apparent potassium intoxication in marine teleosts: a primary toxic effect of copper. Aquaculture 17: 231.

California State Water Resources Control Board (CSWRCB). Procedures manual for toxicity test protocols developed by the Marine Bioassay Project. Report No. 96-1WQ, CSWRCB, Sacramento, CA. January 1996.

Coale KH, Bruland KW (1988) Copper complexation in the Northeast Pacific. Limnol. Oceanogr. 33 (5): 1084-1101.

Corbett JR, Wright K, Baillie AC (1984) The Biochemical Mode of Action of Pesticides. Second Edition. Academic Press. Harcourt Brace Jovanovich, Publishers, London. pp. 9 - 12.

Crespo S, Karnaky KJ (1983) Copper and zinc inhibit chloride transport across the opercular epithelium of seawater-adapted killifish (*Fundulus heteroclitus*). J. Exp. Biol. 102: 337.

DeLisle PF, Roberts MH Jr. (1988) The effect of salinity on cadmium toxicity to the estuarine mysid *Mysidopsis bahia*: role of chemical speciation. Aquat. Toxicol. 12: 357-370.

Engel D, Sunda WG, Fowler BA (1981) Factors affecting trace metal uptake and toxicity to estuarine organisms. I. Environmental Parameters. In: Vernberg FJ, Calabrese A , Thurberg FP, Vernberg WB (eds) Biological Monitoring of Marine Pollutants. Academic Press, New York. P. 144.

Heath, AG (1987) Water Pollution and Fish Physiology. CRC Press, Boca Raton, FL. p 109 -113.

Hemmer MJ, Middaugh DP, Chason V (1992) Acute sensitivity of larval topsmelt, (*Atherinops affinis)*, and inland silversides, (*Menidia beryllina*), to insecticides. Environ. Tox. and Chem. 11(3): 401-408.

Jones MB (1975) Synergistic effects of salinity, temperature and heavy metals on mortality and osmoregulation in marine and estuarine isopods (*Crustacea)* Mar. Biol 30: 13-20.

Lauenstein GG, Cantillo AY (ed.s) (1993) Sampling and Analytical Methods of the National Status and Trends Program National Benthic Surveillance and Mussel Watch Projects 1984 - 1992, Volume III, Comprehensive Descriptions of Elemental Analytical Methods. NOAA Technical Memorandum NOS ORCA 71. National Oceanic and Atmospheric Administration, Silver Spring Maryland pp.105

Levitan WM, Taylor MH (1979) Physiology of salinity-dependent napthalene toxicity in *Fundulus heteroclitus*. J. Fish. Res. Bd. Can. 36: 615-620.

McLusky DS, Hagerman L (1987) The toxicity of chromium, nickel and zinc: effects of salinity and temperature, and the osmoregulatory consequences in the mysid *Praunus flexuosus*. Aquatic Toxicology 10: 225-238.

Middaugh DP, Anderson BS (1994) Utilization of topsmelt, *Atherinops affinis* in environmental toxicology studies along the Pacific coast of the United States. Rev. Environ. Toxicol. 5: 1-50.

Sprague JB (1985) Factors that modify toxicity. In: Rand GR, Petrocelli SR, (ed.s) Fundamentals of Toxicology. Hemisphere Publishing Corp., Washington, DC, pp 136-146.

Stagg RM, Shuttleworth TJ (1982) The accumulation of copper in *Platichthys flesus L.* and its effects on plasma electrolyte concentrations. Fish Biol. 20: 491-500.

Sunda W, Guillard RRL (1976) The relationship between cupric ion activity and the toxicity of copper to phytoplankton. J. Mar. Res. 34: 511-529.

Thurberg FP, Dawson MA, Collier RS (1973) Effects of copper and cadmium on osmoregulation and oxygen consumption in two species of estuarine crabs. Mar. Biol. 23: 171-175.

United States Environmental Protection Agency. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to west coast marine and estuarine organisms (first edition). Chapman, G.A., Denton, D.L, and Lazorchak, J.M. (eds). EPA/600/R-95-136. Office of Research and Development, Cincinnati, OH, 657 pp, 1995.

Voyer RA, Heltsche JF, Kraus RA (1979) Hatching success and larval mortality in an estuarine teleost, *Menidia menidia (Linnaeus)*, exposed to cadmium in constant and fluctuating salinity regimes. Bull. Environ. Contam.Toxicol. 23: 475-481.

Voyer RA, Cardin JA, Heltshe JF, Hoffman GL (1982) Viability of embryos of the winter flounder *Pseudopleuronectes americanus* exposed to mixtures of cadmium and silver in combination with selected fixed salinities. Aquat. Toxicol. 2: 223-233.

Weber CI, Horning WB II, Klemm DJ, Neiheisel TW, Lewis PA, Robinson E. L, Menkedick J, Kessler F (1988) Short-term methods for estimating the chronic toxicity of effluents and receiving waters to marine and estuarine organisms. EPA/600/4-87/028. Final/Technical Report. U.S. Environmental Protection Agency, Cincinnati, OH. pp 417.

Section 2

Mysid (*Holmesimysis costata*) Test Experiments

Precision and sensitivity of the seven-day growth and survival toxicity test

INTRODUCTION

As part of the national effort to improve and maintain water quality, waste water dischargers are required to comply with toxicity objectives in National Pollutant Discharge Elimination System (NPDES) permits [1]. Following guidance from the U.S. Environmental Protection Agency (EPA) [2,3], the California State Water Resources Control Board (SWRCB) has supported the development of toxicity test protocols for use in compliance monitoring of ocean discharges. As part of the SWRCB Marine Bioassay Project, sublethal toxicity test protocols have been developed for mollusc [4,5], algal [6,7], and fish [8,9] species, as well as for the mysid crustacean *Holmesimysis costata*, which is the subject of the present study. The former two protocols are widely used for permit monitoring in California, while the latter two have been recommended for this purpose. All are included in the EPA protocols for west coast species [10]. The present study provides an evaluation of the sensitivity, precision, and logistical feasibility of the seven-day growth and survival toxicity test using *H. costata*.

H. costata is ecologically important as a food web link between lower trophic levels (including detritus, primary producers, and zooplankters) and higher predators, especially juveniles of economically important kelp forest fishes [11]. *H. costata* is easily collected year-round from the surface canopy of *Macrocystis* giant kelp forests, and gravid females can be isolated in the laboratory to release the mysid juveniles used as test organisms. Juvenile mysids can be produced from laboratory populations, but additional research is necessary to sustain high fecundity rates in culture [12]. *H. costata* ranges from La Jolla, California to the Queen Charlotte Islands, British Columbia [13], and its biology has been described from field and laboratory studies [12,14]. Many previous studies have referred to this organism as *Acanthomysis sculpta* (Tattersall), though this designation has been revised by Holmquist [13,15] in her comprehensive review of northeast Pacific mysids. We have considered her designation of *Holmesimysis costata* to be definitive.

Mysid crustaceans are often used in toxicity investigations. Mysids are sensitive to numerous toxicants [16,17,18,19], and previous studies have shown *H. costata* to be a

sensitive species [20,21,22,23]. Although at least one long-term toxicity experiment has been described with this species [24], most previous studies with *H. costata* have utilized (96–h) mortality tests [20,21,25,26,27,28]. A number of acute toxicity test protocols have been published [26,29,30], but estimation of chronic toxicity for permit monitoring requires protocols measuring sublethal endpoints on sensitive life stages [31]. Sublethal endpoints often include test organism growth, because this endpoint represents an integration of a number of physiological processes and is easily quantified [31,32]. The present study evaluates a 7-d *H. costata* growth and survival protocol to determine its suitability as a short-term indicator of chronic toxicity. Experiments are described that estimate intralaboratory precision based on reference toxicant testing and side-by-side effluent tests, and interlaboratory precision based on testing of split samples of whole complex effluent at different laboratories. Results from testing with reference toxicants were compared to results reported from other standardized protocols to evaluate test sensitivity. The degree to which the 7-d test estimated chronic toxicity was investigated through comparison with a 24-d exposure to the same reference toxicant. Measurement of mysid length and weight were compared to determine the most appropriate method for quantifying the growth endpoint.

METHODS

Facilities

Experiments were conducted at the California Department of Fish and Game Marine Pollution Studies Laboratory (MPSL) at Granite Canyon. The laboratory is located on the Big Sur Coast in Monterey County, California, with a seawater intake remote from known sources of pollution.

Holmesimysis costata (Holmes) 7-d toxicity test protocol

These experiments were conducted according to methods described in the 7-d *H. costata* toxicity test protocols [10,22]. These two protocols are essentially the same, except that the earlier protocol version [22] specifies mysid growth measured as length, whereas the latter version [10] specifies growth measured as weight. Both length and weight were measured in a number of tests in the present study to determine which method provided the more appropriate measure of toxicant impacts on mysid growth, as discussed below.

The test procedure involved capturing adult mysids from the field by sweeping a dipnet beneath the surface canopy fronds of the giant kelp *Macrocystis pyrifera*. Gravid females, identified by their extended gray brood pouches, were isolated in the laboratory and placed in a screened (≈ 1 mm) compartment within a larger aquarium. Juveniles were released during nocturnal molting, and were collected the following day (day 0) by removing the adults (in the screened compartment) and siphoning the juveniles into another container. Juveniles were cultured for three to four days in aerated containers on a diet of *Artemia* nauplii and finely ground Tetramin® flake food. On day three or four, juveniles were isolated randomly into sorting cups, then five juveniles were delivered to each test container via pipet with a minimal amount of residual seawater. Test solution volume was 200 ml in each of five replicate containers for each toxicant dilution. Test conditions were as follows: salinity was 33 ± 2 ppt, photoperiod was 16 hours of light, 8 hours darkness at ambient laboratory light intensity (≈100 lux), and target temperature was either 12° or 15°C, depending on whether mysids were originally collected north or south of Pt. Conception, California. (Temperature effects on toxicant sensitivity have not been investigated, but no significant difference in control growth has been observed in mysids exposed for 7-d at the different temperatures [23].) Test temperatures were allowed to vary by ± 2°C from target temperature. Dissolved oxygen was between 60% and 100% saturation. Test solutions were renewed at 48 and 96-h by replacing 75% of the test solution volume. Salinity, pH, and dissolved oxygen concentration were measured in one random replicate from each test concentration at the beginning and end of the test, and test

temperature was monitored continuously. Experimental evaluation of a number of test parameters has been reported previously [23].

Mysid mortality was recorded daily. Mysids that did not respond to a probe were considered dead, and were removed from the container. Feeding rate, 40 *Artemia* nauplii per mysid per day, in one or two daily rations, was adjusted daily for mortality. At the end of the test, the total number of survivors in each replicate was recorded, dead mysids were removed, and surviving mysids were prepared for measurement of length, weight, or both. Mysids to be weighed were in poured onto a fine mesh screen (≈ 200 μm). The screen was dipped in distilled water to remove salt, and mysids were placed with forceps into prenumbered and preweighed microweighboats (≈4 mg) and dried (to constant weight) for at least 24-h at 60° C. After cooling in a dessicator, weighboats were weighed to the nearest microgram. Weighboat tareweight was subtracted from the total, and the difference was divided by the number of mysids to get a mean individual mysid weight per replicate. For length measurement, surviving mysids were poured into a test tube, which was then put in a conventional freezer until ice began to form. Mysids were chilled to near freezing to minimize abdominal flexing during preservation with formalin (5%); this allowed mysids to remain relatively straight for measurement. Mysid images projected through a projecting microscope were measured from the base of the eyestalk to the tip of the telson, and their length was quantified by comparison to a similarly projected image of a stage micrometer. For measurement of both length and weight, mysids were processed for length measurement as above, then rinsed in distilled water and processed for weighing, as above.

Growth measurements were made only in those concentrations where survival was not significantly less than in controls. This practice is common in growth tests [31], because organisms that survive high toxicant concentrations are often larger than average (unpublished data), and because the purpose of measuring sublethal endpoints is to determine whether evidence of chronic toxicity can be detected at concentrations lower than those causing acute mortality.

No observed effect concentrations (NOECs) were calculated for both survival and growth data by analysis of variance followed by Dunnett's test. Survival data were

arcsine squareroot transformed to normalize data prior to analysis. Reported LC50 values were calculated using the Trimmed Spearman-Karber Method [33,34]. The concentrations at which growth was inhibited by 10 and 25% (IC10 and IC25 values) were calculated from growth data using the ICp Calculation Program (Release 1.0, 1988, Battelle Columbus in cooperation with US EPA ERL-Duluth).

Intralaboratory precision

Six reference toxicant tests were conducted using zinc sulfate to evaluate intralaboratory variability of the protocol. The range of temporal variability was estimated by conducting three tests consecutively during the month of March, and comparing these results with results from tests conducted in January, April, and October. Zinc concentrations were 0 (control), 10, 18, 32, 56, and 100 µg/L in all tests. All zinc concentrations are reported as nominal values. Chemical verification of zinc concentrations in the 10 to 100 µg/L range from previous studies at MPSL have shown nominal and measured concentrations to vary by a CV of 14% [21]. Two sets of two side-by-side tests were conducted using split samples of complex effluent. Test precision in all intra- and interlaboratory tests was estimated by calculating the coefficient of variation (CV = standard deviation ÷ mean) of test LC50 and ICp values.

Interlaboratory testing with complex effluents

Four interlaboratory comparisons are described here. The first was conducted using reconstituted lyophilized bleached Kraft mill effluent. Effluent was lyophilized (freeze dried), stored frozen, split into samples for each laboratory, shipped overnight on ice, and reconstituted at each laboratory by mixing the powdered sample with dilution water, according to methods described by Higashi *et al.* [35]. Lyophilized effluent is a useful reference material because the toxicity of the original effluent is retained throughout extended frozen storage [35]. Participating laboratories were the Marine Pollution Studies Laboratory at Granite Canyon (MPSL; Laboratory 1) and a university laboratory with experience in toxicity testing with other mysid species (Laboratory 2). Two-day-old juvenile mysids, released from field-collected females at MPSL, were

packed in plastic bags containing oxygenated seawater and shipped overnight to the second laboratory in a cooler with blue ice (approximately 15°C). Juvenile mysids from the same cohort were packaged similarly and held at MPSL for one night prior to the test. The following day, the three-day-old mysids were taken from the shipping bags at each laboratory and randomly assigned to test containers. Test concentrations were 0 (control), 0.5%, 1.0%, 2.0% and 4.0% effluent.

The second, third, and fourth interlaboratory tests were conducted with whole effluent from a large publicly owned treatment works (POTW) that discharges to California coastal waters. A 24-h composite effluent sample was collected at the treatment plant the day before each test, split into identical one-liter samples, and shipped overnight on ice to the participating laboratories. A second sample was collected and shipped 48-h later to accommodate renewals at 48 and 96-h in all tests. Participating in the POTW interlaboratory tests along with MPSL were two private laboratories (3 and 4) with limited previous experience with this test protocol. Two-day-old mysids were shipped to participating laboratories the day before testing, as described above. Effluent test concentrations were 0 (control), 0 (brine control), 1.8%, 3.2%, 5.6%, 10% and 18%. The brine control solution used spring water in place of effluent at the same concentration as in the highest effluent treatment (18%), with enough hyper-saline brine added to adjust the spring water up to 33 ppt; dilution water (laboratory seawater) was then added to fill the liter mixing flask (as was done in mixing the effluent solutions). Hyper-saline brine was made by freezing 1 µm-filtered seawater until brine salinity reached approximately 70 ppt. Tests at all laboratories in both trials were conducted according to the 7-d growth and survival protocol [22]. All participating laboratories used their own dilution water, brine, glassware and laboratory equipment.

Evaluation of the sensitivity of various endpoints to complex effluent

A number of tests were conducted to evaluate the sensitivity of various endpoints to whole effluent samples. Effluent samples from a variety of sources were used, including an advanced primary sewage treatment plant treating waste water from a small agricultural municipality, combined primary/secondary treated effluent from a large

industrial city, and combined primary/secondary treated effluent from an extended industrial urban area. Effluent concentrations varied depending on the expected toxicity of the effluent, but were typically in a series of 0 (control), 0 (brine control), 1.8%, 3.2%, 5.6%, 10% and 18% effluent. In nine effluent tests, survival was recorded after 4 and 7-d. NOEC and LC50 values were compared to evaluate the sensitivity of 96-h ("acute") and 7-d ("short-term chronic") endpoints. NOEC values from growth endpoints were compared to the survival NOECs to determine relative sensitivity.

Chronic toxicity was estimated by exposing juvenile mysids to zinc sulfate in a 24–d static renewal test, the results of which were compared to a 7-d growth and survival test initiated at the same time. Three-day-old mysids from the same cohort were randomly distributed to test containers for both exposures, and the test containers were interspersed randomly within the same water bath. Mysid mortality was recorded daily, and the lengths of surviving mysids were measured at the end of each exposure. In both experiments, six zinc concentrations, 5.6, 10, 18, 32, 56, and 100 µg/L, and a control were each replicated eight times. Test solutions were renewed after days 2 and 6 in both exposures, and again every 96-h in the longer exposure. Water quality parameters were measured at the beginning and end of each test and before renewals on days 6 and 16. Salinity in all treatments was elevated by evaporation at day 6, to as much as 5 ppt above the target salinity of 34 ppt, but this condition was restored to normal with the renewal, and no other water quality variations were observed. No Observed Effect Concentrations and dose-response data were compared to evaluate the degree to which the seven day test estimated chronic toxicity as observed in the 24-d test.

Both length and weight were measured in five effluent toxicity tests to evaluate the two endpoints. The relative sensitivity of each endpoint was determined by comparison of NOEC, IC10, and IC25 values. The number of tests in which a 25% reduction in growth was detected for each endpoint was also determined. The minimum significant difference (MSD) was calculated for each endpoint in each test to determine the level of difference from the control necessary to indicate statistical significance, given the amount of variability inherent in each endpoint. Finally, three tests were conducted in which initial mysid length and weight were measured to allow change in size, as opposed

to total size, to be evaluated as an appropriate endpoint. Initial lengths of 10 randomly selected mysids and weights of ten sets of five randomly selected mysids were measured to determine initial mysid size at the beginning of these tests.

Sodium azide tests

Sodium azide is a cytochrome oxidase inhibitor that was evaluated as a candidate reference toxicant. Sodium azide was considered an appropriate reference toxicant because we anticipated that it might inhibit mysid growth at concentrations lower than those causing significant mortality, therefore providing reference toxicant data useful for evaluating both endpoints. Three tests were conducted, two range finding tests and a definitive test with more closely spaced toxicant concentrations. Nominal sodium azide concentrations in the definitive test were 0 (control), 10, 18, 32, 56, 100, and 320 µg/L. Azide concentrations were not chemically verified, but the compound is readily soluble and extremely stable in seawater [36].

RESULTS

Intralaboratory precision

Test response to zinc was generally consistent, with a 25% CV among test LC50s (Table 2). No Observed Effect Concentrations (NOECs) varied by no more than one dilution concentration in five of six tests. A sixth test (October, Table 2) produced a NOEC two concentrations lower. The dose response curve for this test was similar to that of the others, except that the 18 µg/L treatment produced markedly higher mortality than in other tests (Figure 1). Of the five replicates of this treatment, one had 80% mortality, reducing the treatment mean by approximately 10%. Because NOEC values are noncontinuous variables dependent on the dilution interval, the coefficient of variation (CV) among NOECs was not calculated. Three zinc tests conducted during different weeks within the same month produced very similar LC50 values, with a CV of 5 %. None of the zinc tests induced a growth response at concentrations below those causing

significant mortality. While this result was consistent, the precision of the growth endpoint could not be quantified from the zinc tests.

In side-by-side tests with complex effluent, LC50 values varied by an average of 7% (Table 3). Dose response curves for the first pair of tests were remarkably similar (Curves A1 and A2, Figure 2). Tests with a variety of other effluents showed a similar pattern of variable response at low effluent concentrations, followed by pronounced decreases in survival above NOEC values. The primary treated effluent produced greater toxicity at lower concentrations relative to the combined primary/secondary effluents. Growth data was not available from one of the tests, while the other test did not produce a growth effect.

In a second side-by-side effluent test, IC10 values for growth in length were most highly variable, due to variability in response at low concentrations (Table 3). IC25 values varied by 15% for length and 23% for growth.

Table 2. Intralaboratory precision data from six zinc reference toxicant tests of the 7-d mysid growth and survival protocol.

	Survival		Growth	
Season	NOEC	LC50	NOEC	IC10
January	32	69	≥ 32	nc
March	32	59	≥ 32	nc
March	56	62	≥ 56	nc
March	56	65	≥ 56	nc
April	32	47	≥ 32	nc
October	10	32	≥ 10	nc
Mean ± SD (LC50)		56 ± 13.8		nc
Coefficient of Variation		25%		

No Observed Effect Concentrations (NOECs) and Median Lethal Concentrations (LC50s) are in µg/L zinc.

"nc" indicates that point estimates (IC10s) could not be calculated from growth data in the zinc tests.

Table 3. Intralaboratory precision data from two trials in which split effluent samples were tested concurrently in the same laboratory.

	Survival (7-d)		Growth (Length)			Growth (Weight)		
Test	NOEC	LC50	NOEC	IC10	IC25	NOEC	IC10	IC25
A1	5.6	7.7	na	na	na	na	na	na
A2	5.6	8.3	5.6	na	na	na	na	na
A Intralab CV:		5%						
B1	10	10.7	10	2.6	6.8	5.6	1.0	5.6
B2	10	12.1	< 1.8	0.8	5.5	5.6	1.2	7.8
B Intralab CVs:		9%		75%	15%		13%	23%

na = data not available.

CV = Coefficient of Variation,

IC10, IC25 = Inhibition Concentrations estimating 10 and 25 percent growth inhibition.

Figure 1. Dose-response curves from mysid zinc tests conducted at different times of year.

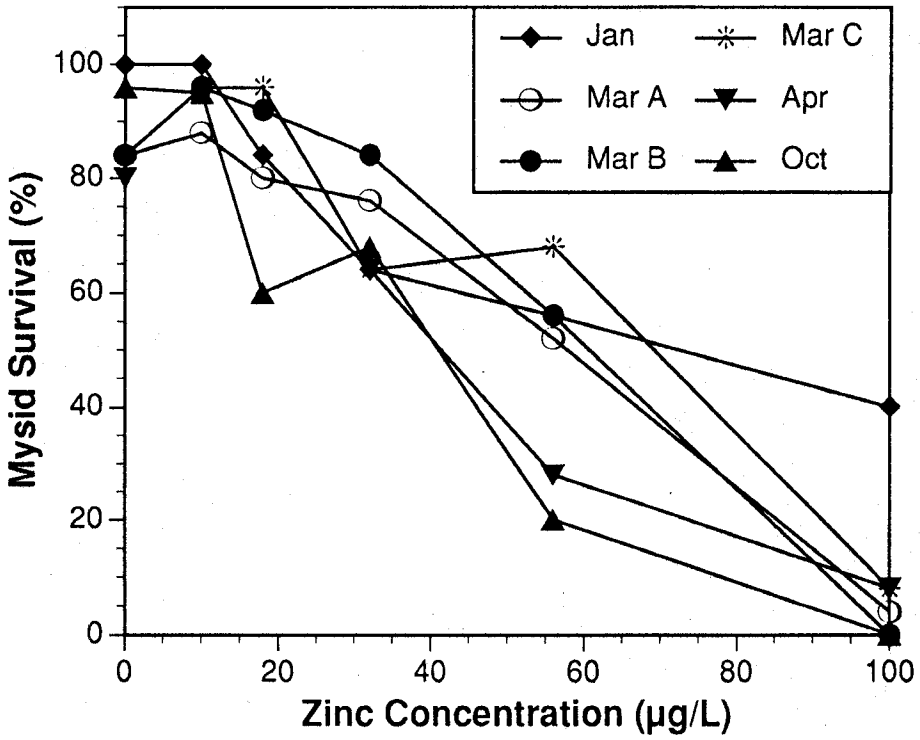
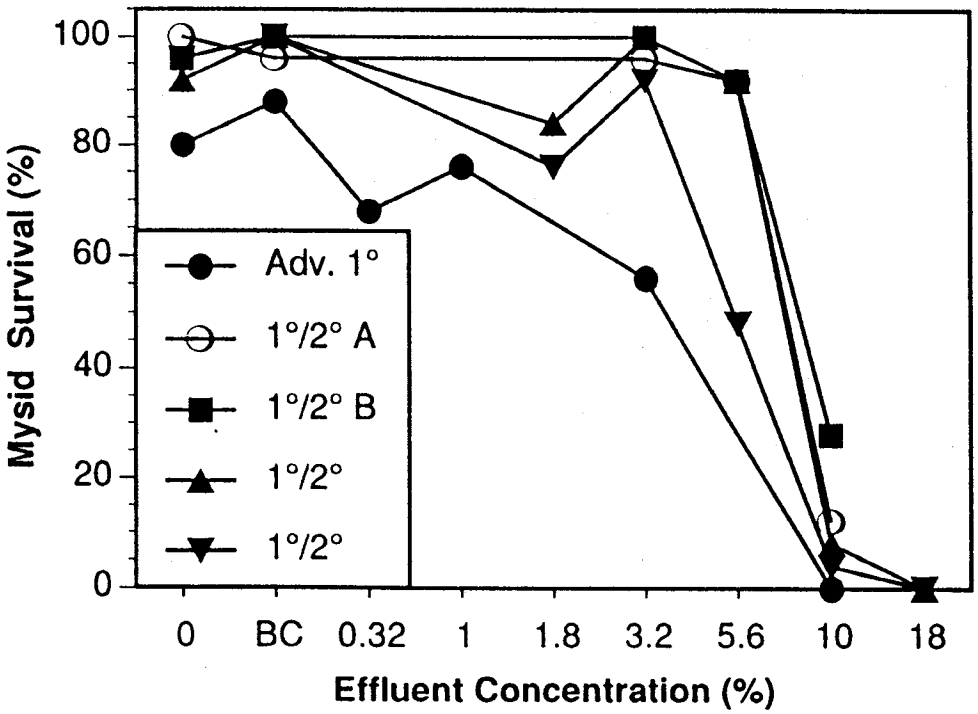


Figure 2. Mysid test response to effluents from primary and mixed primary/secondary treatment plants. A and B represent split samples of the same effluent.



Interlaboratory precision

Survival NOEC values were the same for both laboratories in three of the four interlaboratory tests, with the NOECs varying by one concentration in the fourth test. The interlaboratory CV among LC50s ranged from 7% to 30%, with a mean CV of 14%. In the three interlaboratory tests where the growth endpoint was measured at both laboratories, the NOECs were the same in one test, both higher than the mortality NOEC in a second test, and offset by one dilution concentration in the third test (Table 4). Growth was a more sensitive indicator of effluent toxicity than survival in three of the four interlaboratory tests. Variability among IC10 values for growth measurements were relatively high, again due to variation in response at low concentrations. Variation among IC25 values was lower, with interlaboratory CVs of 8% for length in one set of tests and 21% for weight in a second set.

Table 4. Interlaboratory precision data for the 7-d *Holmesimysis costata* growth and survival toxicity test protocol.

			Growth			Survival	
Test	Effluent	Laboratory	NOEC	IC10	IC25	NOEC	LC50
C1	BKME	1	0.5^{L}	0.74	1.25	1.0	2.0
C2	BKME	2	0.5^{L}	0.54	1.12	1.0	1.8
	CV			22%	8%		7%
D1	POTW	1	$\geq 3.2^{L}$	2.80	nc	3.2	5.1
D2	POTW	3	$\geq 3.2^{L}$	nc	nc	3.2	4.1
	CV			nc	nc		14%
E1	POTW	1	5.6^{W}	4.5	8.3	10.0	11.7
E2	POTW	4	na	na	na	10.0	12.8
	CV			nc	nc		6%
F1	POTW	1	3.2^{W}	1.26	7.34	5.6	9.7
F2	POTW	4	5.6^{W}	7.34	9.94	10.0	14.9
	CV			100%	21%		30%
	Mean Interlaboratory Cvs:			61%	15%		14%

^{L,W} Growth in length was measured in tests 1 and 2, weight was measured in tests 3 and 4.

NOEC, ICp, and LC50 values are in units of percent effluent.

CV is the coefficient of variation (%) for test point estimate values.

BKME is bleached Kraft mill effluent.

POTW is effluent from publicly owned treatment works.

Sensitivity of various endpoints to complex effluent

LC50 values for 96 h data were higher those for the 7-d data in all nine tests evaluated (Table 5), indicating increased sensitivity in the 7-d test. Ninety-five percent confidence intervals around 4-d and 7-d LC50 values did not overlap in six of the nine tests, indicating significant differences. Two of the nine tests had higher 4-d NOEC values than 7-d NOECs. In six tests, the NOEC values were the same, and in one test, the 7-d NOEC was higher, due to increased control mortality near the end of the test. Growth NOEC values were lower than survival NOEC values in seven of the nine tests.

Comparison of the 7-d protocol with a 24-d exposure indicated that this increased exposure significantly increased mysid response to zinc (Figure 3). While survival in controls was not different for the two exposure periods, survival values in all toxicant concentrations above the long-term NOEC of 5.6 µg/L were significantly lower in the long-term test than in the 7-d test (ANOVA, $p < .0001$). After the first 7 d of the 24-d test, mysid survival was similar to that in the 7-d test, with an LC50 of 60 µg/L compared with 50 µg/L in the 7-d test. The LC50 for survival after the 24-d exposure was 7.8 µg/L. Growth was not significantly inhibited at concentrations less than or equal to the NOEC in either test.

Table 5. Comparison of the sensitivity of different protocol endpoints in effluent tests. Four-day LC50 and NOEC values were derived from observations of mysid survival after the first 96-h of the 7-d tests. Control survival was recorded at the end of each 7-d test.

		Control Survival (7-d)		LC50		NOEC (Survival)		NOEC (7-d Growth)	
Test	(Effluent)	Dilution	Brine	4-d	7-d	4-d	7-d	Length	Weight
G	(R)	80%	88%	4.6	4.1	3.2	3.2	1.0	na
A2	(M)	96%	100%	9.0	8.3	5.6	5.6	≥ 5.6	na
H	(M)	92%	100%	14.6*	7.6*	10	5.6	1.8	na
I	(M)	96%	100%	7.8*	5.4*	3.2	3.2	< 1.8	na
B1	(U)	96%	88%	12.4	10.7	10	10	3.2	5.6
B2	(U)	96%	88%	15.5*	12.1*	5.6	10	< 1.8	5.6
E	(U)	100%	84%	17.6*	11.7*	10	10	3.2	5.6
J	(U)	96%	88%	5.1*	4.5*	3.2	3.2	< 1.8	1.8
K	(U)	96%	100%	14.3*	12.0*	10	5.6	≥10.0	5.6
Mean		94%	93%	11.2	8.5	6.8	6.3	4.3-	4.8

R Effluents are from a Rural municipality with a primary treatment plant.

M Effluents are from a large Municipal area with a combined primary/secondary treatment plant.

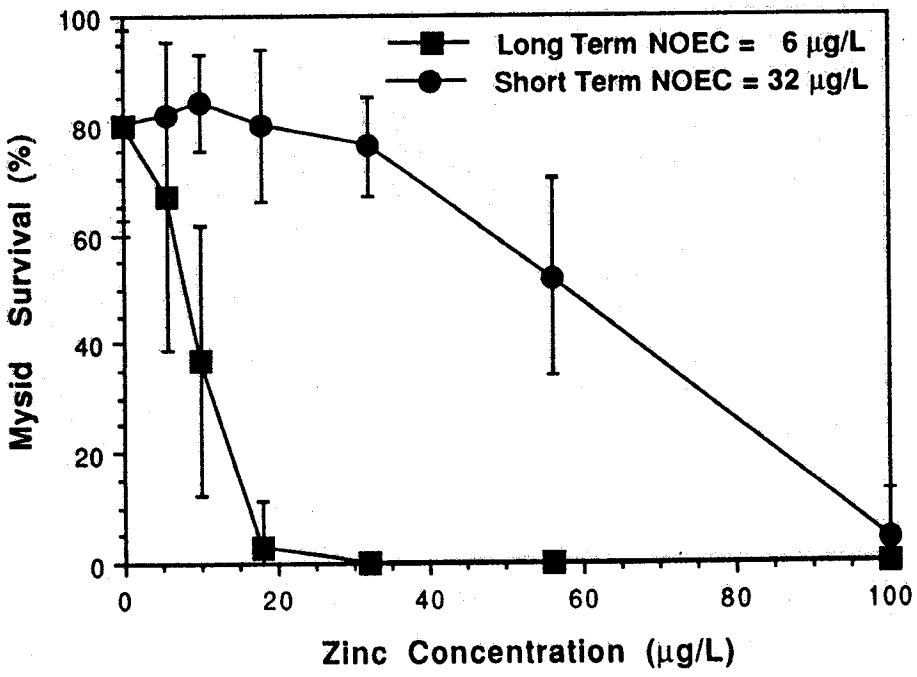
U Effluents are from an extended Urban area with a combined primary/secondary treatment plant.

No Observed Effect Concentrations (NOECs) and Median Lethal Concentrations (LC50s) are in units of percent effluent.

* Asterisks indicate that 95% confidence intervals from 4-d and 7-d LC50s did not overlap.

- NOEC values for length (n = 9) were averaged without regard to ≤ or ≥ signs.

Figure 3. Comparison of long- and short-term mysid tests with zinc.



Evaluation of length and weight endpoints

Weight and length data were compared in five effluent tests. Three of these tests measured initial mysid size, allowing calculation of ICp values for both total final size and change in size over the test period. Both IC10 and IC25 values were lower for weight data than for length data in four of five tests in which they could be compared, indicating greater sensitivity of the weight endpoint. This comparison included tests for which the ICp values could not be calculated, since the inability to calculate the value is indicative of decreased sensitivity. In one test, ICp values for change in length were lower than for change in weight.

NOEC values differed among the two endpoints in all tests, with weight NOECs generally one concentration lower than survival NOECs (Tables 5 and 6), while length NOECs were more variable. Variability among replicates was lower in the length data, leading to lower MSD values. These lower MSDs allow smaller differences from control values to be statistically significant, even when point estimates indicate a lower overall effect level.

Table 6. Comparison of growth endpoints.

	Growth (Length)				Growth (Weight)			
Test	NOEC	MSD%	IC10	IC25	NOEC	MSD%	IC10	IC25
K	≥10	4.3	nc	nc	5.6	23.3	2.4	6.2
E1	3.2	8.3	5.4	nc	5.6	17.7	4.5	8.3
B1	10	6.6	7.4	nc	5.6	30.1	1.3	7.2
*B1 (init)	10	(17.9)	2.6	6.8	5.6	(40.4)	1.0	5.6
B2	< 1.8	6.1	5.7	nc	5.6	22.1	1.6	nc
*B2 (init)	< 1.8	(15.7)	0.8	5.5	5.6	(27.5)	1.2	7.8
J	< 1.8	3.0	4.8	nc	1.8	13.0	2.3	nc
*J (init)	< 1.8	(7.7)	2.3	4.8	1.8	(25.5)	1.5	2.7
Mean MSD		5.7				21.2		

CV = Coefficient of Variation, MSD = Minimum Significant Difference

IC10, IC25 = Inhibition Concentrations estimating 10 and 25 percent growth inhibition.

nc = ICp value could not be calculated because no treatment caused a 25% reduction from control values (or 10% reduction in test A). * indicates that control values had initial weights subtracted.

Sodium azide toxicity

Mysids were particularly sensitive to sodium azide NaN_3 (Table 7). However, contrary to our expectations, sodium azide did not significantly inhibit mysid growth at concentrations below those causing significant mortality. Two preliminary tests with azide found 100% mortality in the low mg/L and mid to high µg/L range. A third test with concentrations more closely spaced in the 10 to 320 µg/L range resulted in a NOEC of 100 µg/L and an LC50 of 149 µg/L.

Table 7. Comparison of toxicant sensitivity among toxicity tests using various species. All values are in µg/L.

Tax. Group	Species	Toxicant	Duration	Endpoint	EC50	Source
Crustacean	*H. costata*	Zinc	7 d	Survival	58	[§]
"	*H. costata*	Zinc	7 d	Survival	47	[22]
"	*H. costata*	Zinc	96 h	Survival	97	[21]
"	*Mysidopsis intii*	Zinc	96 h	Survival	331	[42]
"	*Mysidopsis bahia*	Zinc	96 h	Survival	499	[19]
"	*Cancer magister*	Zinc	96 h	Survival	456	[43]
"	*Tisbe holothuridae*	Zinc	48 h	Survival	713	[44]
"	*Callianassa sp.*	Zinc	96 h	Survival	10,200	[45]
Mollusc	*Haliotis rufescens*	Zinc	48 h	Larval Dev.	68	[4]
Crustacean	*H. costata*	Copper	96 h	Survival	17	[21]
"	*Mysidopsis bahia*	Copper	96 h	Survival	181	[19]
Mollusc	*Haliotis rufescens*	Copper	48 h	Larval Dev.	9	[46]
Crustacean	*H. costata*	Mercury	72 h	Survival	9	[20]
"	*Mysidopsis bahia*	Mercury	72 h	Survival	4	[20]
"	*H. costata*	DSS	72 h	Survival	960	[20]
"	*Mysidopsis bahia*	DSS	72 h	Survival	3,800	[20]
"	*H. costata*	PCB	72 h	Survival	12.5	[20]
"	*Mysidopsis bahia*	PCB	72 h	Survival	14.2	[20]
Crustacean	*H. costata*	Endosulfan	96 h	Survival	0.2	[47]
Mollusc	*Crassostrea gigas*	Endosulfan	48 h	Larval Dev.	55	[48]
"	*Crassostrea virginica*	Endosulfan	48 h	Larval Dev.	460	[49]
Crustacean	*H. costata*	NaPCP	96 h	Survival	84	[47]
Mollusc	*Crassostrea gigas*	NaPCP	48 h	Larval Dev.	48	[50]
"	*Crassostrea virginica*	NaPCP	48 h	Larval Dev.	370	[51]
Crustacean	*H. costata*	NaN_3	7 d	Survival	149	[§]
Algae	*Macrocystis pyrifera*	NaN_3	48 h	Germination	79,000	[23]
Fish	*Atherinops affinis*	NaN_3	7 d	Survival	49,000	[52]
Mollusc	*Mytilus californianus*	NaN_3	48 h	Larval Dev.	28,000	[36]

NaPCP is sodium pentachlorophenate; NaN3 is sodium azide;
DSS is dodecyl sodium sulfate; PCB is polychlorinated biphenyl (Aroclor 1254).
Lower EC50 (= LC50) values indicate greater sensitivity to the toxicant.
§ indicates data is from the present study.

Test Success Rate

Data from all tests reported met the protocol test acceptability criteria, including 80% survival in controls (Table 5). Of 40 tests conducted at MPSL and other laboratories over a four-year period of protocol development, using both effluents and reference toxicants, 80% of tests met the control survival criterion in dilution water controls, and 75% of the tests met this criterion in both dilution water and brine controls [23]. All tests conducted at MPSL during interlaboratory testing met test acceptibility criteria. Four of seven tests conducted at remote laboratories met all criteria, with all test failures occurring during the laboratories' initial trial of the protocol. The level of effort involved in conducting the test was similar to that required for other 7-d tests commonly used for regulatory applications [31].

DISCUSSION

Toxicity testing is widely used for effluent discharge permit monitoring because it provides a direct measurement of the integrated effects of all effluent constituents on representative aquatic organisms. However, in order for specific test methods to be effective, their reliability must be demonstrated, and they must be capable of detecting the presence of toxic contaminants at levels low enough to protect the greatest number of species inhabiting receiving waters. While toxicity tests using marine organisms from the east coast of the United States have been accepted by the EPA and are currently used throughout the nation [31], tests using west coast species are only recently being considered for adoption on a national level [10], based on recent data indicating the relative precision and sensitivity of these toxicity tests. The mysid *Holmesimysis costata* has been the subject of experimentation in an effort to develop a short-term crustacean toxicity test to estimate the chronic toxicity of effluents to marine life. The results of these efforts provide an indication of the suitability of the 7-d growth and survival protocol for routine monitoring of effluent discharges.

Estimation of test precision

Evaluating precision is the primary means of determining the reliability of a toxicity test method for use as a regulatory tool [37]. Toxicity test accuracy cannot be measured because the response of a test organism cannot be compared to a standard or reference value [31]. Precision, however, can be determined by estimating variability among test results, both from reference toxicant tests conducted at different times (intralaboratory temporal variability), or from tests of split samples conducted either side-by-side at one laboratory or concurrently at multiple laboratories (spatial or interlaboratory variability).

Intralaboratory precision

Six tests using the reference toxicant zinc sulfate conducted at different times in the same laboratory produced a CV of 25% among test LC50s (Table 2). Single chemical toxicity test CV values from the literature range from 2 to 94 %, with a mean of 29 % for numerous studies on a variety of organisms (Table 8). Agreement among three *H. costata* zinc tests conducted during the same month was much closer, with a CV of 5 %. This difference in CV values may indicate seasonal variation in the sensitivity of the test; tests conducted in Spring produced consistently lower LC50 values than the test conducted in October. A similar pattern was indicated by previous data from 96-h tests using this species [21]. Though not sufficiently investigated in toxicity testing, seasonal variability does exemplify potential differences between chemical and biological monitoring. There is no apparent basis for seasonal variation in chemical analyses, but it is reasonable to expect that seasonal factors affecting organisms and ecosystems in the field might modify toxic response [7]. Seasonal changes in sensitivity to toxicants may reflect natural conditions in receiving waters, and should be studied further before being categorized as random variation in toxicity test results.

Intralaboratory precision was also estimated from results of two sets of concurrent tests conducted at the same laboratory using split samples of complex effluent. These tests generated identical survival NOECs, and CVs of 5 and 9% among test LC50s,

perhaps providing a reasonable estimate of minimum test variability. Intralaboratory CV values found in the literature from numerous laboratory studies with complex effluent ranged from 0 to 135 % with a mean of 21% (Table 8). Intralaboratory variability among growth endpoints in a split sample effluent test was greater than that for survival data, with the greatest variability between IC10 values from length data. IC25 values may be a better indicator of dose response because of variation in response at low effluent concentrations (Figure 3), and because of the greater biological significance associated with a higher degree of growth inhibition.

Table 8a. Intralaboratory precision comparisons among toxicity test protocols.

Organism	Source	Endpoint	CV Single Chemical		CV Effluent	
			Mean (Range)	N[a]	Mean (Range)	N[a]
H. costata	[§]	7 d Survival	25%	1 (6)	7% (5 - 9)	1 (2)
H. costata	[§]	7 d Length	NA	NA	15%	1 (2)
H. costata	[§]	7 d Weight	NA	NA	23%	1 (2)
M. bahia	[38]	7 d Survival	24% (13 - 36)	4 (2 - 5)	NA	NA
H. costata	[21]	4 d Survival	23%	1 (3)	NA	NA
M. bahia	[38]	2 - 4 d Surv.	14% (4 - 24)	3 (3 - 6)	38% (6 - 135)	9 (2 - 6)
Various[b]	[41]	short-term[b]	54% (36 - 68)	5 (5)	43% (19 - 69)[c]	5 (3 - 5)
Various[d]	[38]	acute[d]	33% (3 - 94)	22 (2 - 14)	17% (0 - 135)	46 (1 - 12)
Various[d]	[38]	chronic[d]	32% (2 - 83)	36 (2 - 15)	7% (0 - 20)	8 (2 -4)
Mean			29%		21%	

NA: not available

Table 8b. Interlaboratory precision comparisons among toxicity test protocols.

Organism	Source	Endpoint	CV Single Chemical		CV Effluent	
			Mean (Range)	N[a]	Mean (Range)	N[a]
H. costata	[§]	7 d Survival	NA	NA	14% (6 - 30)	4 (2)
H. costata	[§]	7 d Length	NA	NA	8%	1 (2)
H. costata	[§]	7 d Weight	NA	NA	21%	1 (2)
H. costata	[21]	4 d Survival	13% (1 - 24)	2 (2)	NA	NA
M. bahia	[38]	4 d Survival	41% (18 - 61)	8 (2 - 10)	65% (18 - 166)	6 (2 - 12)
Various[b]	[41]	short-term[b]	66% (44 - 110)	7 (5)	58% (33 - 88)	4 (3 - 5)
Various[d]	[38]	acute[d]	46% (18 - 143)	22 (2 - 10)	34% (0 - 166)	24 (2 - 16)
Various[d]	[38]	chronic[d]	38% (7 - 71)	14 (5 - 10)	34% (0 - 83)	10 (3 - 9)
Mean			41%		33%	

CV is the coefficient of variation, based on LC50s for survival data and IC25s for growth data.

[a] Number of studies, with tests per study in parentheses.

[b] Protocols included short-term larval development tests with bivalves, fertilization tests with echinoderms, and algal reproductive tests.

[c] Effluent was lyophilized.

[d] Protocols include tests with mysids, daphnids, fish, echinoderms, and algae.

NA: not available

Interlaboratory tests

The interlaboratory precision data for *H. costata* survival is indicated by CV values ranging from 7 to 30%, with a mean of 14%, lower than reported mean CV values for numerous tests and laboratories (Table 8). These tests were designed to estimate precision that might be expected under routine effluent testing conditions [39]. While juvenile mysids for interlaboratory tests in this study were provided from the same source, no other factors were controlled. Laboratories used different dilution water, glassware, personnel and facilities.

Interlaboratory precision of the growth endpoint was indicated by a mean CV of 61% for IC10s and 15% for IC25s. As above, the IC25 data provided the greater precision. No comparison data was found for interlaboratory variability from 7-d growth tests. However, available data on interlaboratory variation in effluent tests using sublethal endpoints indicates CVs range from 33 to 88%, with a mean of 58% (Table 8 [41]).

Intra- and interlaboratory CV values for the *H. costata* toxicity test were comparable to CV values for intra- and interlaboratory precision estimates for analytical chemistry techniques used in permit monitoring. DeGraeve [37] summarized EPA intralaboratory chemical precision data [40], finding mean CV values from replicate laboratory spikes to be: 37 % for semi-volatiles using GC-MS (range 13 to 90%); 30 % for pesticides and PCB mixtures using GC with electron-capture detector (range 20 to 61%); and 36 % for PAHs measured by GC with a flame ionization detector (range 20 to 50 %). Analysis of 14 metals with inductively coupled plasma-atomic emission spectroscopy generated a mean interlaboratory CV value of 13 % (range 1.1 to 45 %).

Test feasibility

The *H. costata* test required a level of effort similar to that required by other commonly used 7-d growth and survival tests for mysids and fish larvae [31]. During four years of test development, 75% of the 7-d tests met all test acceptability criteria and were considered successful (n = 40). Of the interlaboratory tests conducted, four of seven tests met acceptability criteria at both participating laboratories, while the remaining three interlaboratory tests were successful at one of the two participating laboratories (11 of 14 overall, or 79% success rate). In all cases, test failures occurred during the laboratories' first trial of the protocol. These success rates compare with rates of 67 to 77% for a large interlaboratory study involving five laboratories, five test species, and three toxicants [41]. Test organism supply also dictates feasibility. A number of commercial suppliers exist to provide *H. costata* juveniles, but reliable large-scale supply must await increased demand by testing laboratories.

Test endpoints

The 7-d *H. costata* test was significantly less sensitive to zinc than was a 24-d exposure (Figure 3). The longer-term test produced zinc survival NOEC and LC50 values of 5.6 and 7.8 µg/L, much lower than the respective values of 32 and 50 µg/L for the 7-d test. Survival data recorded at day 7 of the long-term test produced an LC50 of 60 µg/L. The similarity of this value with the 7-d test LC50 (50 µg/L) indicated that differences in 24-d and 7-d test results were due to exposure duration, rather than to initial mysid condition or laboratory artifacts. While growth was not significantly inhibited even in the longer-term zinc test, it was the most sensitive endpoint in seven of nine (78%) of the effluent tests reported here, including interlaboratory tests (Tables 4 and 5).

The 7-d test endpoints were slightly more sensitive than was the survival response measured after 96 h of exposure. The 7-d LC50 values were lower than the 96-h LC50 values in each of nine effluent tests evaluated (Table 5), and significantly lower in six of those tests. The 7-d NOEC values (from either the growth or survival endpoints) were lower than the 96-h survival NOEC values in two of the nine effluent tests, but late

mortality in controls caused the 4-d NOEC to be lower than the 7-d in one test. Though the response observed in the 7-d test was only slightly stronger than that observed after 96-h, it should be noted that a sensitive life stage (early juvenile) of a relatively sensitive species was used in both exposures. *H. costata* 96-h juvenile survival tests are generally more sensitive than acute tests using adults of other species [21].

While NOEC values for length data were lower than those for weight data in three of five tests where both endpoints were measured, this apparent sensitivity is primarily the result of low between-replicate variability. This low variability is probably associated with the fact that multiple length measurements (one per surviving mysid) are averaged within each replicate, while there is only a single weight measurement per replicate. The MSD values for length data averaged 8.7%, while weight MSDs averaged 25%. Point estimates indicate that effluents had a greater impact on mysid weight than length, as indicated by lower ICp values for weight. Weight NOEC values tended to be similar to weight IC25 values (Table 6), while the association between length NOEC and ICp values was less consistent. Given that test endpoints should measure biologically significant effects, the weight measurement appears more likely to result in NOEC values indicative of reduced growth as opposed to reduced variability.

Sodium azide toxicity

One definitive test with sodium azide indicated that this chemical did not inhibit mysid growth at sublethal concentrations. The established mode of action of this cytochrome oxidase inhibitor is to block electron transport during mitochondrial respiration, a process we predicted would significantly inhibit growth at concentrations below those causing significant mortality. While growth inhibition was not observed, the mysid test did prove to be exceptionally sensitive to azide. The LC50 value for this test was orders of magnitude lower than that for tests using organisms from other phyla (Table 7).

Relative test sensitivity

To further evaluate the relative sensitivity of the 7-d *H. costata* toxicity test, data were compared to available literature values for common toxicants (Table 7). Calculated median effective concentrations (EC50s or LC50s) for the *H. costata* test were found to be lower for zinc, copper, dodecyl sodium sulfate and Aroclor 1254 than for toxicity tests with other crustacean species, but higher for mercury relative to a similar test with the east coast mysid *Mysidopsis bahia*. The *H. costata* test was generally less sensitive to trace metals than were tests with molluscan species. Its sensitivity to the insecticide endosulfan was greater than for tests with other phyla, but it was less sensitive to the organic herbicide sodium pentachlorophenate than were tests reported for molluscan species.

Given the limited knowledge available on the effects of pollutants to the great diversity of marine species and life stages, it is essential to adopt relatively sensitive toxicity tests to provide regulatory protection for potentially impacted organisms. From the data that are available, it is clear that toxicant sensitivity is species- and protocol-specific, and that impacts to local biological communities will be most safely estimated by employing a variety of species, including those likely to inhabit receiving waters. The protocol for *H. costata* provides a high level of sensitivity to a number of toxicants (Table 7), and its level of precision and logistical feasibility make it suitable for use in evaluating potential pollutant impacts. Available data suggest that *H. costata* test precision is comparable to toxicological and chemical methods currently used in effluent monitoring. Toxicity tests with crustacean species are essential to effective monitoring programs where effluents may contain insecticide wastes, and the natural distribution of *H. costata* makes it an appropriate crustacean species for determining the toxicity of effluent discharges in west coast marine waters.

REFERENCES

1. **U.S. Environmental Protection Agency**. 1984. Development of water quality-based permit limitations for toxic pollutants: National policy. Fed. Reg. 49:9016-9019.

2. **U.S. Environmental Protection Agency**. 1985. Technical support document for water quality-based toxics control. Office of Water, Washington, DC.

3. **U.S. Environmental Protection Agency. 1991**. Technical support document for water quality-based toxics control. EPA/505/2-90-001, PB91-127415. Office of Water, Washington, DC.

4. **Hunt, J.W. and B.S. Anderson.** 1989. Sublethal effects of zinc and municipal effluents on larvae of the red abalone *Haliotis rufescens*. *Marine Biology* **101**:545-552.

5. **Hunt, J. W. and B.S. Anderson**. 1993. From research to routine: A review of toxicity testing with marine molluscs. In W.G. Landis, J.S. Hughes, and M.A. Lewis, eds., *Environmental Toxicology and Risk Assessment*. ASTM STP 1179, American Society for Testing and Materials, Philadelphia, PA, pp. 320-339.

6. **Anderson, B.S. and J.W. Hunt**. 1988. Bioassay methods for evaluating the toxicity of biocides, heavy metals, and sewage effluent using microscopic stages of giant kelp *Macrocystis pyrifera* (Agardh) : A preliminary report. *Mar. Environ. Res.* **26**: 113 - 134.

7 **Anderson, B.S., J.W. Hunt, S.L. Turpen, A.R. Coulon, M. Martin**. 1990. Copper toxicity to microscopic stages of giant kelp *Macrocystis pyrifera*: interpopulation comparisons and temporal variability. *Mar. Ecol. Prog. Series* **68**: 147 - 156.

8. **Anderson, B.S., D.P. Middaugh, J.W. Hunt, and S.L. Turpen**. 1991. Copper toxicity to sperm, embryos, and larvae of topsmelt *Atherinops affinis*, with notes on induced spawning. *Mar. Environ. Res.* **31**: 17-35.

9. **Anderson, B.S., J.W. Hunt, H.R. McNulty, S.L. Turpen, and M.Martin**. 1994. Off-season spawning and factors influencing toxicity test development with topsmelt *Atherinops affinis*. *Environ. Toxicol. and Chem.* **13**:479-486.

10. **Chapman, G. A., D. L. Denton, and J. M. Lazorchak.** 1995. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to west coast marine and estuarine organisms. First Edition. EPA/600/R-95-136. US Environmental Protection Agency, Cincinnati, Ohio.

11. **Hobson, E. S. and J. R. Chess**. 1976. Trophic interactions among fishes and zooplankters near shore at Santa Catalina Island, California. *Fish Bull.* **74**: 567-598.

12. **Turpen, S.L., J.W. Hunt, B.S. Anderson, and J.S. Pearse**. 1994. Population structure, growth, and fecundity of the kelp forest mysid *Holmesimysis costata* in Monterey Bay, California. *J. Crustacean Biol.* **14**:657-664.

13. **Holmquist, C**. 1979. *Mysis costata* Holmes, 1900, and its relations (Crustacea, Mysidacea). *Zool. Jb. Syst.* **106**: 471 - 499.

14. **Green, J. M.** 1970. Observations on the behavior and larval development of *Acanthomysis sculpta* (Tattersall), (Mysidacea). *Can. J. Zool.* **48**:289-292.

15. **Holmquist, C**. 1981. The genus *Acanthomysis* Czerniavsky, 1882 (Crustacea, Mysidacea). *Zool. Jb. Syst.* **108**:386-415.

16. **Nimmo, D. R., L. H. Bahner, R. A. Rigby, J. M. Sheppard and A. J. Wilson Jr.** 1977. *Mysidopsis bahia*: An estuarine species suitable for life-cycle toxicity tests to determine the effects of a pollutant. ASTM STP 634, American Society for Testing and Materials, Philadelphia, PA, pp. 109-116.

17. **Gentile, S. M., J. H.Gentile, J. Walker, and J. F. Heltshe**. 1982. Chronic effects of cadmium on two species of mysid shrimp: *Mysidopsis bahia* and *Mysidopsis bigelowi. Hydrobiologia* **93**:195-204.

18. **Breteler, R. J., J. W. Williams and R. L. Buhl**. 1982. Measurement of chronic toxicity using the oppossum shrimp *Mysidopsis bahia. Hydrobiologia* **93**: 189-194.

19. **Lussier, S. M., J. H. Gentile and J. Walker.** 1985. Acute and chronic effects of heavy metals and cyanide on *Mysidopsis bahia* (Crustacea: Mysidacea). *Aquat. Toxicol.* **7**: 25-35.

20. **Tatem, H.E. and A.S. Portzer**. 1985. Culture and toxicity tests using Los Angeles district bioassay animals, *Acanthomysis* and *Neanthes*. Misc. Paper EL-85-6, U.S. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, MS.

21. **Martin, M., J.W. Hunt, B.S. Anderson, S.L. Turpen and F.H. Palmer**. 1989. Experimental evaluation of the mysid *Holmesimysis costata* as a test organism for effluent toxicity testing. *Environ. Toxicol. and Chem.* **8**: 1003 - 1012.

22. **Hunt, J.W., B.S. Anderson, S.L. Turpen, H.R. Barber, M. Martin, D.L. McKeown, and F.H. Palmer.** 1991. Marine Bioassay Project sixth report: Interlaboratory comparisons and protocol development with four marine species. Report # 91-21WQ. California State Water Resources Control Board, Sacramento, CA.

23. **Anderson, B.S., J.W. Hunt, H.R. McNulty, M.D. Stephenson, F.H. Palmer, D.L. Denton, M. Reeve**. 1994. Marine Bioassay Project seventh report. Division of Water Quality report no. 94-2WQ. State Water Resources Control Board, Sacramento, CA.

24. **Davidson, B. M., A. O. Valkirs and P. F. Seligman**. 1986. Acute and chronic effects of tributyltin on the mysid *Acanthomysis sculpta* (Crustacea, Mysidacea). Proceedings, Oceans 86, Washington, D.C., Sept 23-24, 1986, pp. 1219-1225.

25. **Machuzak, M. J. and T. K. Mikel, Jr.** 1987. Drilling fluid bioassays using Pacific ocean mysid shrimp, *Acanthomysis sculpta*, a preliminary introduction. In W. J. Adams, G. A. Chapman, and W. G. Landis, eds., Aquatic Toxicology and Hazard Assessment: 10, ASTM STP 971. American Society for Testing and Materials, Philadelphia, PA, pp. 447-453.

26. **Reish, D.J. and J.A. LeMay**. 1988. Bioassay manual for dredged sediments. U.S. Army Corps of Engineers, Los Angeles District, Los Angeles, CA. 160 pp.

27. **Asato, S.L. and D.J. Reish**. 1989. The effects of heavy metals on the survival and feeding of *Holmesimysis costata* (Crustacea: Mysidacea). Biologia Marina, Memorias del VII Simposium, La Paz, Baja California Sur, Mexico. pp. 113 - 120.

28. **Singer, M.M., D.L. Smalheer and R.S. Tjeerdema**. 1990. Toxicity of an oil dispersant to the early life stages of four California marine species. *Environ. Toxicol. Chem.* **9**:1387-1395.

29. **Anderson, B.S., J.W. Hunt, S.L Turpen, A.R. Coulon, M. Martin, and F.H. Palmer.** 1990. Procedures manual for conducting toxicity tests developed by the Marine Bioassay Project. Report No. 90-10WQ. State Water Resources Control Board, Sacramento, CA.

30. **ASTM**. 1993. Standard guide for conducting static and flow-through acute toxicity tests with mysids from the west coast of the United States. Designation E 1463 - 92; 1993 Annual Book of Standards, Volume 11.04. American Society for Testing and Materials, PA, pp. 1278 - 1299.

31. **Weber, C.I., W.B. Horning, D.J. Klemm, T.W. Neiheisel, P.A. Lewis, E.L. Robinson, J. Menkedick, and F. Kessler**. 1988. Short-term methods for estimating chronic toxicity of effluents and receiving waters to marine and estuarine organisms. EPA/600/4-87/028. U.S. Environmental Protection Agency, Cincinnati, OH.

32. **Bayne, B.L., M.N. Moore, J. Widdows, D.R. Livingston, and P. Salkeld**. 1979. Measurement of the responses to environmental stress and pollution: studies with bivalve molluscs. *Phil. Trans. R. Soc. Lond. B.* **286**:563-581.

33. **Hamilton, M.A., R.C. Russo, and R.V. Thurston**. 1977. Trimmed Spearman-Karber method for estimating median lethal concentrations in toxicity bioassays. *Environ. Sci. Technol.* **11**: 714-719.

34. **Hamilton, M.A., R.C. Russo, and R.V. Thurston**. 1978. Correction. Environ. Sci. Technol. **12**: 417.

35. **Higashi, R.M., G.N. Cherr, J.M. Shenker, J.M. Macdonald, and D.G. Crosby**. 1992. A polar high molecular mass constituent of Bleached Kraft Mill Effluent is toxic to marine organisms. *Environ. Toxicol. Chem.* **26**:2413-2420.

36. **Cherr, G.N., J. Shoffner-McGee, and J.M. Shenker**. 1990. Methods for assessing fertilization and embryonic/larval development in toxicity tests using the California mussel (*Mytilus californianus*). *Environ. Toxicol. Chem.* **9**:1137 - 1145.

37. **DeGraeve, G.M., J.D. Cooney, D.O. McIntyre, T.L. Pollock, N.G. Reichenbach, J.H. Dean and M.D. Marcus**. 1991. Variability in the performance of the seven-day fathead minnow (Pimephales promelas) larval survival and growth test: an intra- and interlaboratory study. *Environ. Toxicol. Chem.* **10**:1189-1203.

38. **Parkhurst, B.R., W. Warren-Hicks, and L.E. Noel**. 1992. Performance characteristics of effluent toxicity tests: summarization and evaluation of data. *Environ. Toxicol. Chem.* **11**:771-791.

39. **Anderson, S.L. and T.J. Norberg-King**. 1991. Precision of short-term chronic toxicity tests in the real world. *Environ. Toxicol. Chem.* **10**:143-145.

40. **U.S. Environmental Protection Agency**. 1986. Test methods for evaluating solid waste. SW-846. Office of Solid Waste and Emergency Response, Washington, DC.

41. **BSAB (Bioscience Monitoring Advisory Board).** 1994. West coast marine species chronic protocol variability study: evaluation of results and recommended test methods. BSAB Report #2 for the Washington Department of Ecology, Olympia, WA.

42. **Langdon, C.J., W.K Seim, R.L. Hoffman, L. Weber, and G.A. Chapman**. 1989. Development of protocols for chronic toxicity testing of Pacific marine species. Final Report for EPA cooperative agreement # CR813131, Hatfield Marine Science Center, Newport, OR.

43. **Martin, M., K. Osborn, P. Billig and N. Glickstein**. 1981. Toxicities of ten metals to *Crassostrea gigas* and *Mytilus edulis* embryos and *Cancer magister* larvae. *Mar. Poll. Bull.* 12: 305-308.

44. **Verriopoulos, G. and D. Hardouvelis**. 1988. Effects of sublethal concentration of zinc on survival and fertility in four successive generations of *Tisbe*. *Mar. Poll. Bull.* 19: 162-166.

45. **Ahsanullah, M., D. S. Negilski and M. C. Mobley**. 1981. Toxicity of zinc, cadmium and copper to the shrimp *Callianassa australiensis*. I. Effects of individual metals. *Mar. Biol.* **64**: 299-304.

46. **Hunt, J.W., B.S. Anderson, S.L. Turpen, A.R. Coulon, M. Martin, F.H. Palmer and J.J. Janik**. 1989. Marine Bioassay Project fourth report: Experimental evaluation of effluent toxicity testing protocols with giant kelp, mysids, red abalone, and topsmelt. Report 89 -5WQ. California State Water Resources Control Board, Sacramento, CA.

47. **Martin, M., J.W. Hunt, B.S. Anderson, L. Espinosa and F.H. Palmer.** 1986. Marine Bioassay Project second report: acute toxicity tests and protocols for mysids, red abalone and giant kelp. State Water Resources Control Board, Sacramento, CA.

48. **Dinnel, P. A**. 1991. Toxicity testing with oyster and mussel embryos. In: G.A. Chapman, ed., Culture and toxicity testing of west coast marine organisms, US EPA Newport, OR.

49. **Mayer, F. L. Jr.** 1987. Acute toxicity handbook of chemicals to estuarine organisms. EPA600/8-87/017, U.S. Environmental Protection Agency, Gulf Breeze, FL.

50. **Woelke, C. E**. 1972. Development of a receiving water quality bioassay criterion based on the 48-h Pacific oyster (*Crassostrea gigas*) embryo. Washington Department of Fisheries Technical Report No. 9, Olympia, WA.

51. **Zaroogian**, G.E. 1981. Interlaboratory comparison - acute toxicity tests using the 48-hr. oyster embryo-larval assay. Unpublished EPA Report. Contribution No. 223 of Environmental Research Laboratory, U.S. Environmental Protection Agency, Narragansett, RI.

52. **Anderson, B.S., J.W. Hunt, W.J. Piekarski, B.M. Phillips, M.A. Englund, R.S. Tjeerdema, and J.D. Goetzl.** *In Press.* Influence of salinity on copper and azide toxicity to larval topsmelt *Atherinops affinis* (Ayres). Arch. Environ. Contam. Toxicol.

Section 3

Abalone (*Haliotis rufescens*) Test Experiments

Evaluation of the Use of Small-Volume Containers for Toxicity Tests with Mollusc Larvae

INTRODUCTION

In recent years, 48-h embryo/larval toxicity tests with marine molluscs, including abalone, have been conducted increasingly in small test containers with five to ten mL of test solution (Table 9). This technical development has a number of advantages over conducting tests in large beakers, including greatly reduced sample requirements, logistical benefits, and, most importantly, the elimination of subsampling error and bias in the analysis of toxic effects. In this section, we summarize data available from the literature indicating the level of precision obtained in multiple tests conducted in small volume containers, and examine the comparability of results obtained in large and small test containers. We also present original data from experiments in which we compared the results of 48-h abalone tests conducted in beakers with 200 mL of test solution, as specified in current protocols, and in glass scintillation vials with 10 mL of test solution.

Table 9. Species and applications for small volume toxicity testing with marine molluscs. Test solution volumes are in mL.

Species	Volume	Application	Source
Oyster (*Crassostrea gigas*)	7	pore water	Flegal et al. 1994
Oyster (*Crassostrea gigas*)	10	complex effluents	BSAB 1994, Pastorok et al. 1994
Oyster (*Crassostrea gigas*)	10	pore water	Anderson et al. 1995
Mussel (*Mytilus edulis*)	10	pore water (TIE)	S. R. Hansen & Assoc.(unpubl.)
Mussel (*Mytilus edulis*)	10	complex effluents	BSAB 1994, Pastorok et al. 1994
Mussel (*Mytilus edulis*)	10	sediment elutriate	SFEI 1994, Kline et al. 1993
Mussel (*Mytilus edulis*)	10	pore water	Hunt et al. *In Press*
Mussel (*Mytilus californianus*)	3	sodium azide	Cherr, et al. 1990
Abalone (*Haliotis rufescens*)	10	pore water	Anderson et al. 1993 Sapuder et al. 1994
Clam (*Mulinia lateralis*)	10	effluents (TIE)	Burgess et al. (Draft) Pelletier et al. 1994
Clam (*Mulinia lateralis*)	10	metals, phenol	Morrison & Petrocelli 1990

Small Volume Test Precision

The Washington State Department of Ecology has sponsored large interlaboratory studies to evaluate and determine the precision of a number of toxicity test methods under consideration for NPDES permit monitoring. Among the tests evaluated were the embryo/larval test using both the oyster *Crassostrea gigas* and the mussel *Mytilus spp.* in 30 ml glass vials with 10 ml of test solution. The Washington study found that *Mytilus* tests conducted by five laboratories during seven test periods were successfully completed within quality assurance guidelines 96% of the time; the success rate for tests with *C. gigas* was 81%. Test precision, as indicated by coefficients of variation (CV = standard deviation ÷ mean) were 21% and 29% for interlaboratory effluent tests with *Mytilus* and *C. gigas*, respectively. Intralaboratory precision estimates were similar: 19% for *Mytilus* and 25% for *C. gigas* (BSAB 1994, Pastorok et al. 1994). These tests used lyophilized bleached Kraft mill effluent, which allowed split samples to be reconstituted after extended storage without appreciable loss of effluent toxicity. Tests using cadmium as a reference toxicant produced results that were more variable than those from the lyophilized effluent tests, with interlaboratory CVs of 44% and 59% and intralaboratory CVs of 36% and 54% for *Mytilus* and *C. gigas*, respectively. A similar level of intralaboratory precision (CV = 31%) was obtained from thirteen *Mytilus* embryo-larval tests with cadmium chloride conducted at our laboratory over a two year period using 10 ml volumes in glass scintillation vials. Intralaboratory precion of the abalone embryo/larval test in small volumes (10mL) can be estimated from eleven zinc tests conducted at MPSL, which produced a CV of 13%. These precision estimates compare favorably with published estimates for the precision of other established toxicological and chemical methods (Rue et al. 1988, Morrison et al. 1989, Parkhurst et al. 1992, Hunt and Anderson 1993), and indicate an acceptable level of performance for embryo-larval tests in small volumes.

Evaluation of the Effects of Test Chamber Size

Cherr et al. (1990) exposed *Mytilus californianus* embryos to sodium azide in large and small containers in concurrent tests. The large containers were 600 ml beakers holding 400 ml of test solution, while the small containers were Lab-Tek® (Registered Trademark of Nunc, Inc., Naperville, IL.) mini polystyrene chambers adhered to glass microscope slides, each holding 3 ml of test solution. Three paired tests produced a mean EC50 of 25.5 mg/L in large volumes and 26.2 mg/L in small volumes, indicating that no significant artifacts were introduced by the small containers. At MPSL, we have conducted a number of toxicity tests with red abalone (*Haliotis rufescens*) embryos in large (200 ml) and small (10 ml) containers, and have seen a similar level of agreement. Eleven tests in 10 ml volumes produced a mean zinc EC50 of 44.5 ± 5.7 µg/L, while eight tests in 200 ml volumes produced a mean zinc EC50 of 43.8 ± 9.0 µg/L, again indicating the comparability of results from tests conducted in large and small containers.

Described below are further experiments to investigate possible effects of container size, using two additional toxicants, copper chloride and the organic biocide sodium pentachlorophenate. Both of these compounds are toxic to red abalone embryos at relatively low concentrations (Hunt and Anderson, 1993) that could be affected by differential sorption to container surfaces. Because small volume containers have larger surface area to volume ratios, we investigated whether small containers might adsorb a proportionately greater toxicant mass, resulting in lower test solution concentrations and reduced toxicity.

METHODS

Experimental Design

Two experiments were conducted, one each with copper chloride and sodium pentachlorophenate. Each experiment consisted of four concurrent red abalone 48-h embryo-larval toxicity tests. Each of the four tests was conducted in one of the following types of test container: 600 ml borosilicate glass beakers with 200 ml of test solution, 30 ml borosilicate glass scintillation vials with 10 ml of test solution, 250 ml polypropylene plastic beakers with 200 ml of test solution, or 30 ml high-density polyethylene plastic vials with 10 ml of test

solution. The ratios of wet surface area (cm^2) to test solution volume (ml) for the respective containers were 2.21, 5.42, 1.68, and 4.84, giving the large beakers about two-and-a-half times more surface area per volume than the small vials. There were five replicates of each container treatment at each toxicant concentration.

Each test container was inoculated with an equal density of abalone embryos (10 embryos/ml). Nominal test concentrations for all container types were 0, 2, 4, 8, 16, and 32 µg/L copper, and 0, 10, 20, 40, 80, and 160 µg/L pentachlorophenate. Test solutions at each concentration were delivered to all container treatments from the same volumetric mixing flasks. Copper test solutions were measured at the beginning and end of the exposure period, and from the container surfaces at the end of the experiment, as described below.

Test Protocol

The red abalone embryo-larval test followed a protocol (Anderson et al, 1990) similar to that used in many bivalve tests (for example ASTM 1993). Adult abalone were induced to spawn by exposure to dilute hydrogen peroxide in Tris buffered seawater (Morse et al. 1977). Males and females were spawned in separate containers to isolate eggs and sperm, which were then mixed at appropriate densities. The resulting embryos were rinsed, concentrated into a beaker, and delivered into test containers to provide a final embryo density of 10/ml. Embryos were exposed to static test solutions (without renewal) for 48 h. Water quality parameters of dissolved oxygen, pH, temperature and salinity were measured in each concentration at the beginning and end of each test to verify compliance with quality assurance criteria.

All test containers were cleaned prior to testing using the following procedure: triple rinse with tap water, triple rinse with hexane solvent, triple rinse with deionized water, 24-h soak in 2N hydrochloric acid, triple rinse with deionized water, 24-h soak in deionized water, triple rinse with distilled water, dry in a clean drying oven.

During the 48-h exposure period, the embryos developed into trochophore and then veliger larvae. At the end of the exposure, larvae were prepared for observation in one of three ways. Larvae in the large glass or plastic beakers were poured onto a 37 µm mesh screen, then rinsed into scintillation vials where they were preserved in 3% buffered formalin.

Larvae in small polyethylene vials were poured directly into glass scintillation vials, where they were similarly preserved. Larvae tested in glass scintillation vials were not transferred, but were preserved by addition of formalin directly into the vials. An inverted compound microscope was used to analyze all larvae through the glass bottom of the scintillation vials. We analyzed all of the approximately 100 larvae from each small test container, and a random subsample of 100 larvae from each large container. Larval shell development was the test endpoint measured, and data were reported as normal larvae as a percentage of the total number counted.

Sampling and Analysis of Copper Concentrations

At the beginning and end of the experiment, copper concentrations were measured in test solutions from all four test container types at the three highest nominal concentrations (8, 16, and 32 µg/L). Test solutions with nominal concentrations of 0, 2, and 4 µg/L were not measured because we had less confidence in the reliability of calculations based on measurements nearer the method detection limit (0.4 µg/L). Copper concentrations in acid rinsates of test container surfaces were also measured at these test concentrations. Pentachlorophenate concentrations were not measured because of logistical and funding constraints.

Initial copper concentrations were measured in samples taken directly from the mixing flask that was used to deliver test solution to the toxicity test containers. Final copper concentrations were measured in samples taken from additional replicates (with embryos) of each container type at the three concentrations measured (8, 16, and 32 µg/L). These samples were preserved in high-density polyethylene vials with 1% by volume 14 N quartz-distilled nitric acid.

The following procedure was used to measure copper adsorbed to test container surfaces during testing. At the end of the test period, the entire contents of a test container from each treatment at each concentration was discarded, and the test container was manually centrifuged in a sling for two minutes to remove any remaining test solution. Centrifuge speeds ranged from 90 to 120 rpm for large beakers and 120 to 150 rpm for small vials. All containers were visually inspected for consistency of droplet removal. Estimates of the

remaining water volume ranged from 0.1 to 0.3 ml per container. Test containers were then rinsed by swirling 2N nitric acid to contact all inner surfaces at least three times. Ten ml of acid was used to rinse the small containers, and 30 ml was used to rinse large containers. The acid rinsate was then poured into high-density polyethylene vials for chemical analysis. All copper concentrations were analyzed using a Perkin Elmer Zeeman/3030 graphite furnace atomic absorption spectrophotometer with a method detection limit of 0.4 µg/L. All measured concentrations were corrected by subtraction of appropriate matrix blanks. Analytical precision, as indicated by replicate measurements of 32 µg/L samples, was 3.8% (mean = 31.9, s.d = 1.2 µg/L). Total copper was measured, and there was no attempt to distinguish "free" copper ions from complexed forms.

Data Analysis

Toxicity data were used to compare the effects observed in tests with each of the four different container types. EC50 data for each toxicant and test container type were calculated using the trimmed Spearman-Karber method (Hamilton et al. 1977). A two-way analysis of variance (ANOVA) was used to determine the statistical significance of differences between concentrations and container types. Proportion data (percent normally developed larvae) were normalized by transformation to the arcsine square root prior to performing the analysis of variance (Zar 1984).

Chemistry data were used to calculate changes in copper concentration during the exposure period, and to determine the mass of copper removed from container surfaces after testing. The calculated reduction in test solution copper concentration (µg/L) due to adsorption to container surfaces was calculated by dividing the concentration of copper in the acid rinsate (µg/L) by the volume of acid (L) to get the mass of copper from the container surface (µg), and then dividing this mass by the original test solution volume (L). This value was then compared to changes between initial and final copper concentrations as measured directly from the test solutions.

RESULTS AND DISCUSSION

Effects of Container Type on Toxicant Concentration

Measured copper concentrations were unchanged during the 48-h experiment in 2 of the 12 test solutions sampled. In 6 of the 12 samples, concentrations decreased (by an average of 21%: Table 10). Copper was detected in all but one of the acid solutions used to rinse container surfaces after the test (n = 12). Calculations indicated that copper recovered in the rinsate accounted for 4 to 27% of the initial copper in test solutions. This mass of copper is one to two orders of magnitude greater than could be accounted for by test solution droplets left behind on container surfaces after centrifugation. Of those samples where test solution chemistry indicated a decline in concentration over the test period, copper recovered from container surfaces was sufficient to account for 15 to 59% of the decrease (Table 10). In the 4 remaining samples, measured copper concentrations increased (by an average of 14%), either through contamination or measurement error.

Decreases in test solution copper tended to be affected more by container material than container size, with more copper recovered from plastic surfaces than from glass. Large and small glass containers had strikingly similar levels of container surface copper on a mass per test solution volume basis, and there was similar agreement between values for large and small plastic containers (Table 10).

Table 10. Surface area to volume ratios and test EC50's.

Container Type	Surface Area to Volume Ratio	Copper Chloride EC50	(C.I.)	Sodium Pentachlorophenate EC50	(C.I.)
Plastic Vials	4.84	9.4	(8.3 - 10.6)	nd	nd
Plastic Beakers	1.68	13.4	(11.7 - 15.3)	54.3	(47.6 - 61.8)
Glass Vials	5.42	9.5	(8.4 - 10.8)	49.8	(45.5 - 54.4)
Glass Beakers	2.21	10.6	(9.9 - 11.5)	65.0	(57.2 - 73.9)

C.I. is 95% confidence interval for EC50 estimates.

Effects of Container Type on Toxicity

There were significant differences in embryo-larval development among tests conducted in different types of containers (ANOVA: $p < .0001$). Contributing to this statistical significance was the low between-replicate variability characteristic of the test, with a mean between-replicate CV of 10%. The differences among tests are indicated by the EC50 values for the respective dose-response curves (Table 11, Figures 4 and 5). In all cases, tests conducted in small vials resulted in lower EC50s than tests in beakers, indicating greater toxicity measured in small containers. In two of three available comparisons, there was no overlap of EC50 95% confidence intervals from large and small container tests. Data were only available for three of the four pentachlorophenate container treatments, since the test conducted in plastic vials produced 100% abnormally developed larvae at all concentrations, including the control. Unexplained uniform container contamination was the probable cause of observed effects in this treatment.

The observed differences in toxicity associated with different container types did not appear to be the result of differential loss of toxicant to container surfaces. The results were counter to the original hypothesis that toxicity would be decreased in smaller containers due to their greater surface to volume ratio and proportionately greater sorption of toxicant from test solutions. In both the trace metal and organic experiments, greater toxicity was observed in the smaller containers (Table 11).

In the experiment with copper, the greatest difference in response was at 16 µg/L, where about 50% of the larvae developed abnormally in the large plastic beakers, as compared with 100% abnormal larvae at that concentration in the other three container treatments (Figure 4). Chemical measurements indicate little difference in final test solution copper concentrations between these treatments (Table 10). All water quality measurements were within normal ranges. Dissolved oxygen averaged 89% saturation, with low variability among measurements. Test solution ammonia concentrations were not measured, but embryo densities were the same in all containers.

In the pentachlorophenate experiment, the dose-response curve was steepest for the test in small glass vials, with 100% abnormal larvae at the 80 µg/L concentration, as compared to about 60% abnormal larvae at that concentration in the other containers (Figure 5). There were more abnormal larvae in the glass vial controls compared to the other tests, but larval response in the low pentachlorophenate concentrations was similar to that observed in the other container treatments. Organic chemical concentrations were not measured, so there are no chemical data with which to evaluate the observed differences in effects. Measured water quality parameters were all within normal ranges.

In both experiments there were differences in larval development between test container treatments, but these differences did not appear to be the result of toxicant sorption related to container size. Further testing would be necessary to investigate alternative causes for observed differences.

Table 11. Measured copper concentrations from test solutions and rinsates of container surfaces.

Nominal Concentration (µg/L)	Initial Cu Conc. (µg/L)	Final Cu Conc. (µg/L)	Rinsate Cu Conc. (µg/L acid)	Rinsate Cu per Test Sol. Vol. (µg/L)	Sorbed Cu* as a % of: Initial Cu Conc	Sorbed Cu* as a % of: Cu lost over Exp period
Glass Vial						
8	6.4	7.4	nd	nd	nc	nc
16	12.0	13.0	1.0	1.0	8	nc
32	33.0	27.0	1.5	1.5	5	25
Glass Beaker						
8	6.4	2.6	3.4	0.6	9	15
16	12.0	15.0	5.1	0.8	7	nc
32	33.0	28.0	8.2	1.4	4	27
Plastic Vial						
8	6.4	6.9	1.6	1.6	25	nc
16	12.0	12.0	3.2	3.2	27	nc
32	33.0	28.0	2.5	2.5	8	50
Plastic Beaker						
8	6.4	6.3	8.8	1.4	23	nc
16	12.0	12.0	13.6	2.2	19	nc
32	33.0	27.0	21.6	3.6	11	59

* Copper collected in rinsates of container surfaces is assumed to have been adsorbed onto container surfaces from the test solutions.

Method detection limit = 0.4 ppb for all samples
Precision of replicate measurements: CV = 3.8% (n = 5).

Figure 4. Effects of copper on abalone larval development in different types of test containers.

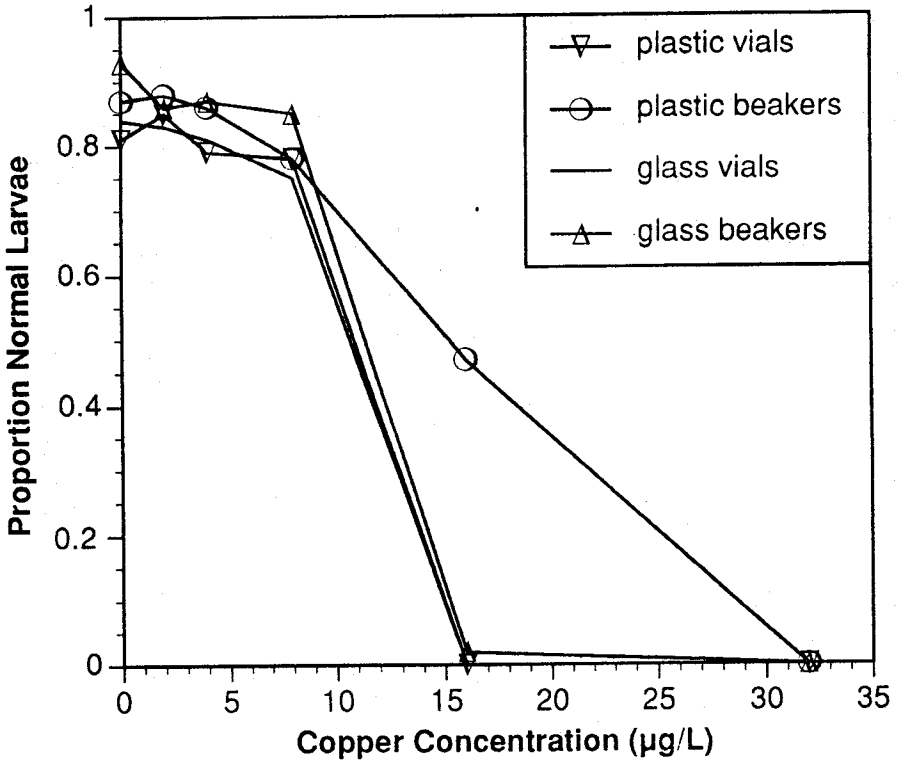
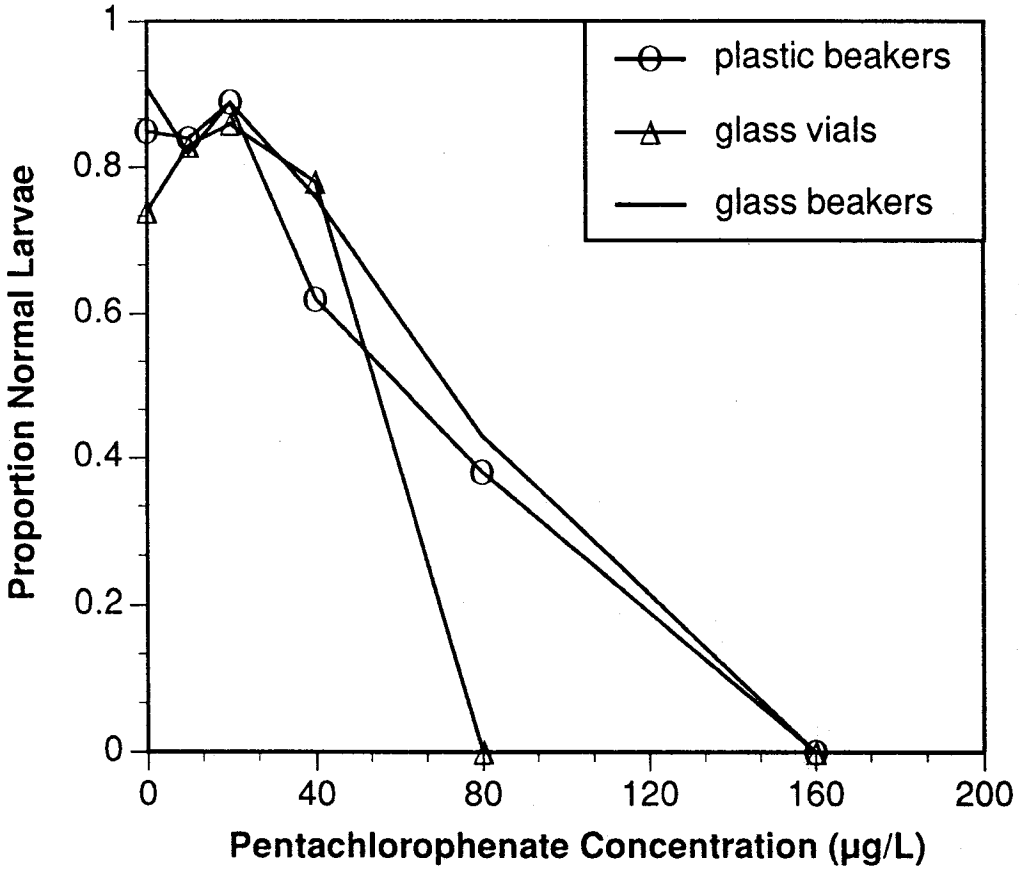


Figure 5. Effects of Sodium Pentachlorophenate on abalone larval development in different types of test containers



SUMMARY

A large number of laboratories throughout the United States and elsewhere are engaged in toxicity testing for permit monitoring activities and other environmental assessments. Many of these toxicity evaluations use a number of molluscan species and protocols, most involving embryo-larval development. The published literature on pollution studies with marine molluscs has been recently reviewed (Calabrese 1984, Hunt and Anderson 1993), but it is likely that the majority of recent data from mollusc testing remains relatively inaccessible in client reports and other formats that are not widely circulated. Recent developments have focused on new applications for existing protocols. Primary among these new applications are assessments of sediment pore water toxicity and identification of individual toxicants in complex mixtures through TIEs. Both of these applications are greatly facilitated through conducting tests in small volumes.

Small volume toxicity tests using mollusc embryos have demonstrated adequate precision for regulatory purposes, as indicated by available intralaboratory data and by the more comprehensive interlaboratory variability study sponsored by the State of Washington (BSAB 1994, Pastorok et al. 1994). Variability among mollusc tests was shown to be lower than that observed in tests with echinoderms and kelp, and compared favorably with published values for other protocols (Rue et al. 1988, Morrison et al. 1989, Parkhurst et al. 1992, Hunt and Anderson 1993).

The side-by-side experiments presented in this chapter indicate that test container type may affect the results of toxicity tests, though the causes of observed differences were unclear. In tests with both trace metal and organic compounds, greater toxicity was observed in smaller containers compared to larger containers with the same toxicant concentrations. This result was contrary to what would be expected if the greater surface area to volume ratio of smaller containers were allowing relatively greater loss of toxicant through sorption to container surfaces. Measured copper concentrations from test solutions and acid rinsates of container surfaces indicated that relative loss of toxicant through sorption was probably influenced more by container material than by configuration (Table 2).

Test solution volume and container type should be considered as possible factors

affecting comparability between results obtained from different exposure systems. This may be especially true of the more hydrophobic compounds. However, data from multiple tests, as presented by Cherr et al. (1990) for marine mussels and in this section for red abalone tests with zinc, indicate that for some toxicants, there may be little appreciable difference in EC50 values from tests conducted in small volumes (3 to 10 ml) as compared to larger volumes (200 to 400 ml).

Since practical laboratory exposure systems are designed for maximum precision, sensitivity and convenience, rather than in an attempt to mimic field conditions, and since there has been minimal previous test container standardization, no advantage is gained by using large test solution volumes. Small volume toxicity testing with marine molluscs eliminates subsampling and handling bias, produces an acceptable level of precision, and increases the feasibility of environmental assessments, especially those involving sediment pore water, TIEs, and other applications where large sample volumes are difficult to obtain.

REFERENCES

Anderson, B. S., Hunt, J. W. and McNulty, H. R., Evaluation of three marine toxicity test protocols for sediment interstitial water, presented at Society of Environmental Toxicology and Chemistry, Houston, November, 1993.

Anderson, B.S., Hunt, J. W., Turpen, S. L., Coulon, A.R., Martin, M., McKeown, D.L., and Palmer, F.H., Procedures manual for conducting toxicity tests developed by the Marine Bioassay Project. Water Resources Control Board, State of California, 90-10WQ, 112, 1990.

Anderson, S. L., Knezovich, J. P., Jelinski, J., Steichen, D. J., The utility of pore water toxicity testing for development of site-specific marine sediment quality objectives for metals, Report LBL-37615 UC-000, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA, 1995.

ASTM, *Annual Book of ASTM Standards*, Vol. 11.04 American Society for Testing and Materials, Philadelphia, PA. 1993.

Bioscience Monitoring Advisory Board, West coast marine species chronic protocol variability study, Washington Department of Ecology, Industrial Section, Olympia, WA, 1994.

Burgess, R.M., K.T. Ho, and G.E. Morrison, Marine toxicity identification evaluation (TIE) guidance document. Phase I. Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C. 20460. Draft.

Calabrese, A., "Ecotoxicological Testing With Marine Molluscs," *Ecotoxicological Testing for the Marine Environment*, Vol. 1, Persoone, G., Jaspers, E., and Claus, C., Eds., State University of Ghent Institute for Marine Scientific Research, Bredene, Belgium, 1984, pp. 455 - 478.

Cherr, G.N., Shoffner-McGee, J., and Shenker, J.M, Methods for assessing fertilization and embryonic/larval development in toxicity tests using the California mussel (*Mytilus Californianus*). Environ. Toxicol. Chem. 9, 1137, 1990.

Flegal, A. R., Riseborough, R. W., Anderson, B.S., Hunt, J.W., Anderson, S., Oliver, J., Stephenson, M., and Packard, R, San Francisco Estuary pilot regional monitoring program: sediment studies, final report, San Francisco Regional Water Quality Control Board, Oakland, CA. July, 1994.

Gary Wortham, S.R. Hansen and Associates, Concord, California, personal communication and unpublished data.

Hamilton, M.A., Russo, R.C. and Thurston, R.V., Trimmed Spearman-Karber method for estimating median lethal concentrations in toxicity bioassay, Environ. Sci. and Technol. 11, 714, 1977.

Hunt, J. W. and Anderson, B.S, From research to routine: a review of toxicity testing with marine molluscs. In *Environmental Toxicology and Risk Assessment. ASTM STP 1179*, Landis, W.G., Hughes, J.S., and Lewis, M.A., eds., Amer. Soc. Test. & Matls., Philadelphia, PA., 1993, 320-339.

Hunt, J.W., B.S. Anderson, and B.M. Phillips. In Press. Small volume toxicity testing with marine molluscs. In P.G. Wells, K. Lee, and C. Blaise (Eds), Microscale aquatic toxicology - Advances, Techniques, Practice. CRC Press, Inc. Boca Raton, FL.

Kline, K. F., Griffen, D. M., and Fisler, M. W., Preliminary TIE assessment of a region-wide toxicity observed in sediments from San Francisco Bay, presented at Society of Environmental Toxicology and Chemistry, Houston, November, 1993.

Morrison, G. E. and Petrocelli, E., Suitability of *Mulinia lateralis* as a euryhaline toxicity test species, presented at 17th Annual Toxicity Workshop, Vancouver, British Columbia, November, 1990.

Morrison, G. E., Torello, E., Comeleo, R., Walsh, R., Kuhn, A., Burgess, R., Tagliabue, M., and Greene, W., Intralaboratory precision of saltwater short-term chronic toxicity tests. Water Pollut. Control Feder. 61, 1707, 1989.

Morse, D.E., Duncan, H., Hooker, N., and Morse, A., Hydrogen peroxide induces spawning in molluscs, with activation of prostaglandin endoperoxide synthetase. Science, 196, 298, 1977.

Parkhurst, B.R., W. Warren-Hicks, and L.E. Noel. Performance characteristics of effluent toxicity tests: summarization and evaluation of data. Environ. Toxicol. Chem. 11, 771, 1992.

Pastorok, R. A., Anderson, J. W., Butcher, M.K. and Sexton, J. E., West coast marine species chronic protocol variability study, Washington Department of Ecology, Industrial Section, Olympia, WA, 1994.

Pelletier, M. C., Kuhn-Hines, A., Burgess, R., Morrison, G., and Ho, K., Utility of the bivalve *Mulinia lateralis* as a marine TIE species, Society of Environmental Toxicology and Chemistry 15th Annual Meeting, Abstract, 1994.

Rue, W. J., Fava, J. A., and Grothe, D. R., A review of inter- and intralaboratory effluent toxicity test method variability, Adams, W. J., Chapman, G. A., and Landis, W. G., eds. *Aquatic Toxicology and Hazard Assessment*: 10th Volume. ASTM Special Technical Publication (STP) 971, Amer. Soc. Test. & Matls., Philadelphia, PA, 1988.

San Francisco Estuary Institute, San Francisco estuary regional monitoring program for trace substances, 1994 annual report, Richmond, CA, 1994.

Sapudar, R.A., Wilson, C.J., Reid, M.L., Long, E.R., Stephenson, M., Puckett, H.M., Fairey, R.A., Hunt, J.W., Anderson, B.S., Holstad, D.R., Newman, J.W., and Birosik, S., Sediment chemistry and toxicity in the vicinity of the Los Angeles and Long Beach Harbors, California State Water Resources Control Board, Sacramento, CA. 1994.

Zar, J.H. *Biostatistical Analysis*, 2nd Edition. Prentice-Hall Inc., New Jersey, 1984.

Section 4

Giant Kelp (*Macrocystis pyrifera*) Experiments

Comparison of short-term and long-term endpoints using copper and sodium azide

INTRODUCTION

It has been demonstrated that reproductive endpoints with Laminarian algae are more sensitive indicators of toxicity than vegetative endpoints (Thompson and Burrows 1984, Chung and Brinkhuis 1986, James et al 1987, Anderson et al 1988, Anderson et al 1990a, see Thursby et al 1993 for review) . However, because of the requirement for longer exposure times and the greater potential for contamination by competing alga species, protocols assessing kelp reproduction have been found to be more difficult to conduct on a routine basis (Anderson et al 1990b). It was concluded that test success rates were considerably higher for the 48-h protocol, and this protocol was recommended for routine use, recognizing that there was a need for some compromise between sensitivity and the requirement for high test success rates for NPDES monitoring programs.

In order to expand on previous research investigating the relationship between short-term and longer-term endpoints, a series of experiments were conducted comparing the relative effects of sodium azide and copper chloride on *Macrocystis* spore germination, gametophyte growth and sporophyte production. Using azide, long-term tests measuring sporophyte production were conducted as either continuous exposures or as 48 and 96-h pulse exposures followed by recovery in uncontaminated seawater. Using copper, long-term tests measuring sporophyte production were conducted as either continuous exposures or as 48-h pulse exposures followed by recovery in uncontaminated seawater. These experiments were designed to determine whether vegetative effects observed in short-term exposures are persistent when cultures are allowed to recover in toxicant-free media, and whether duration of initial exposure affects recovery. Some of these experiments are presented in a recent Marine Bioassay Project publication (Anderson et al 1996).

METHODS

Two separate experiments were conducted using sodium azide. In the first experiment three separate treatments were compared concurrently: 1) a short-term 48-h exposure, 2) a 48-h exposure followed by a 17-day recovery period (48-h-ER), and 3) a 19-d continuous exposure (CE). The second experiment also involved concurrent comparisons of three separate treatments: 1) a short-term 48-h exposure, 2) a short-term 96-h exposure followed by 20-d recovery period (96-h-ER), and 3) a 24-d continuous exposure (CE).

In Experiment 1, germination and germ-tube growth were assessed as described above after a 48-h exposure. For the 48-h exposure-recovery (48-h-ER) treatment, slides containing gametophyte cultures were transferred from azide spiked media to containers with control seawater containing nutrients necessary for gametogenesis [PES (Provasoli 1968) plus germanium dioxide to control diatoms]. In addition, the lighting was changed from cool-white fluorescent to full-spectrum Vitalights® ($50\mu Em^{-2}sec^{-1}$) to induce gametogenesis. Nutrients and full spectrum lighting were also provided to the continuous exposure (CE) treatments after the initial 48 hours, and test solutions for both ER and CE treatments were renewed thereafter every 96 hours. Gametophyte cultures in the ER and CE treatments were allowed to develop for an additional 17 days, when sporophyte numbers were counted.

Procedures in Experiment 2 were identical except for the following modifications: for the 96-h-ER treatments, slides were exposed to azide-free media after an initial 96-h exposure period, instead of after 48 hours. In addition, rather than transferring the slides to new containers for recovery, the azide-spiked media was siphoned out of the orginal ER test containers and replaced with clean sea water.

Procedures for the copper exposures duplicated the first azide experiment described above (Azide - Experiment 1). Germination and germ-tube growth were assessed as described above after a 48-h copper exposure. For the 48-h exposure-

recovery (48-h-ER) treatment, slides containing gametophyte cultures were transferred from copper spiked media to containers with control seawater containing nutrients necessary for gametogenesis [PES (Provasoli 1968) plus germanium dioxide to control diatoms]. The test containers in this case were polyethylene plastic food containers with 200 ml of test solution. As above, the lighting was changed after the initial 48-h from cool-white fluorescent to full-spectrum Vitalights® ($50\mu Em^{-2}sec^{-1}$) to induce gametogenesis. Nutrients and full spectrum lighting were also provided to the continuous exposure (CE) treatments after the initial 48 hours, and test solutions for both ER and CE treatments were renewed thereafter every 96 hours. Gametophyte cultures in the ER and CE treatments were allowed to develop for an additional 20 days, when sporophyte numbers were counted.

Because sporophyte numbers in the copper experiment were counted on two separate microscopes having slightly different magnifications, the relative area covered by each microscopes field of view differed. To account for any effect of this difference in area between microscopes, the sporophyte numbers were normalized to number of sporophytes per unit area (in ml). Since sporophyte numbers in the azide experiments were quantified on the same microscope these were not normalized.

All treatments were analyzed using ANOVA followed by Dunnett's test to determine NOECs. Concentrations causing 50% inhibition of germination and sporophyte production (EC50s) were calculated using the Spearman-Karber method (Hamilton et al 1977). Concentrations causing 50% reduction in germ-tube length were calculated using the EPA ICp program (Norberg-King 1993).

RESULTS

Sodium Azide experiments

As in previous experiments with copper and pentachlorophenate (Anderson and Hunt 1988, Anderson et al. 1990a), the results using sodium azide indicate that the longer-term sporophyte production test was more sensitive than the 48-h germination and growth test. Sporophyte production in the continuous exposure treatment was significantly inhibited ($p = .0001$) at the lowest azide concentration (NOEC <3.2 mg/L; Figure 6a); the NOECs for germination and germ-tube growth were 18 mg/L. Sporophyte production in the 48-h-ER treatments indicates that gametophytes cultured in azide concentrations which significantly inhibit germination and growth are able to recover and undergo reproduction when transferred after 48 h to azide-free media. Although there was greater than 50% reduction in sporophyte numbers at the two highest concentrations (100 and 180 mg/L; Fig. 6a), the effects of these treatments were not statistically significant because of high variability in the data set.

In cultures exposed to azide for 96-h and then transferred to azide-free media, some recovery was indicated, though it is apparent that except in the lowest azide concentration, gametogenesis continued to be suppressed (Fig. 7a). The EC50 for sporophyte production in the 96-h-ER test was 9.2 mg/L, an order of magnitude lower than the EC50 in the 48-h-ER test (128 mg/L; Table 12). The EC50s for the CE tests were comparable in both experiments, approximately 1.0 mg/L. Sporophyte numbers were considerably less in the second experiment, the reason for this disparity is not clear.

Table 12. Results of 48-hour, Continuous exposure (CE) experiments, and 48-h and 96-h Exposure Recovery (ER) experiments with *Macrocystis pyrifera* using sodium azide (in mg/L) as a toxicant. NR indicates no statistical results could be determined.

	Azide Concentrations (mg/L)				
ENDPOINT	**NOEC**	**EC50**	**95% CI**	**IC50**	**95%CI**
Germination (Experiment 1)	18.0	51.53	(49.3-53.8)	96.3	(62.8-208)
Germ tube length (Experiment 1)	18.0				
Sporophyte Production CE (Experiment 1)	<3.2	1.12	(0.9-1.3)		
Sporophyte Production 48-h-ER (Experiment 1)	>180.0	127.7	(28.0-580.8)		
Germination (Experiment 2)	18.0	89.9	(NR)		
Germ tube length (Experiment 2)	10.0			80.1	(NR)
Sporophyte production CE (Experiment 2)	<5.6	1.1	(NR)		
Sporophyte Production 96-h-ER (Experiment 2)	5.6	9.2	(8.6-9.8)		

Figure 6a. Effects of sodium azide on kelp sporophyte production following a 48-h continuous exposure (CE), and after a 48-h exposure followed by a 17-day recovery period (ER).

Figure 6b. Effects of sodium azide on kelp gametophyte germination following a 48-h exposure.

Figure 6c. Effects of sodium azide on kelp gametophyte growth following a 48-h exposure.

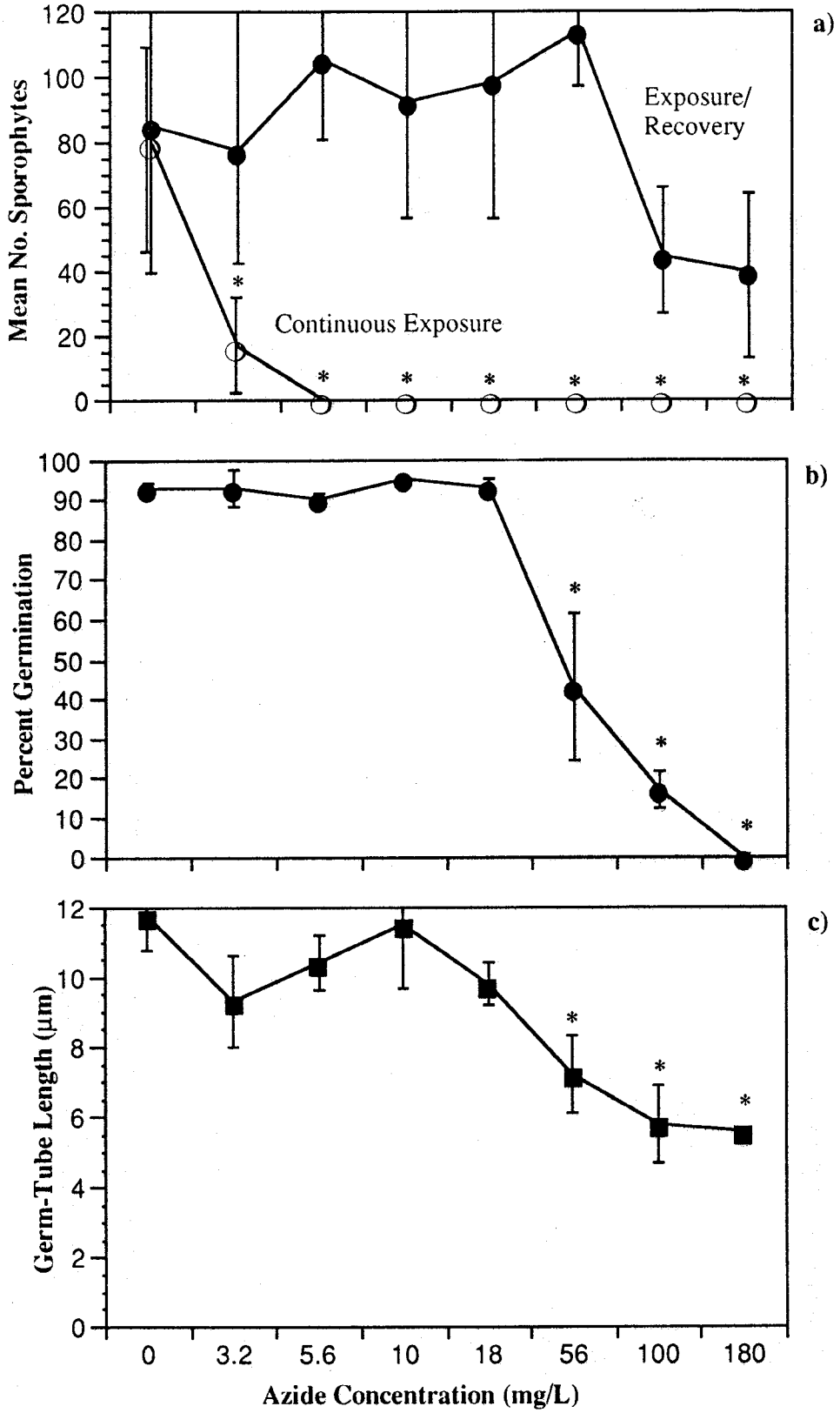
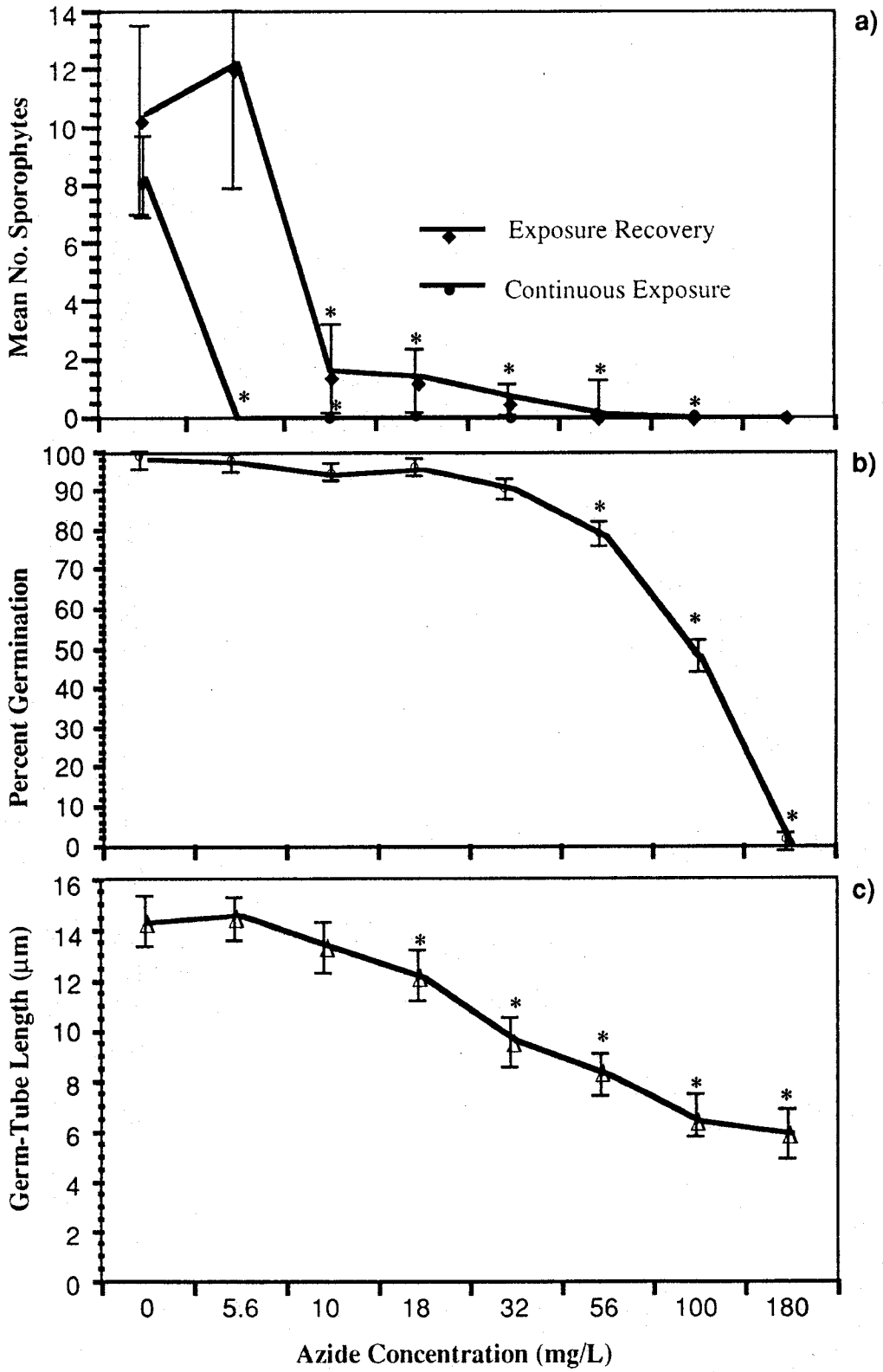

a)
Mean No. Sporophytes
120
100
80
60
40
20
0
Exposure/
Recovery
Continuous Exposure
*
b)
Percent Germination
100
90
80
70
60
50
40
30
20
10
0
c)
Germ-Tube Length (µm)
12
10
8
6
4
2
0
0
3.2
5.6
10
18
56
100
180
Azide Concentration (mg/L)

Figure 7a. Effects of sodium azide on kelp sporophyte production following a 96-hr continuous exposure (CE), and after a 96-hr exposure followed by a 20-day recovery period (ER)

Figure 7b. Effects of sodium azide on kelp gametophyte germination following a 48-hr exposure.

Figure 7c. Effects of sodium azide on kelp gametophyte growth following a 48-hr exposure


a)
Mean No. Sporophytes
Exposure Recovery
Continuous Exposure
b)
Percent Germination
c)
Germ-Tube Length (µm)
0
5.6
10
18
32
56
100
180
Azide Concentration (mg/L)

Copper experiment

Results of the copper test are summarized in Table 13. NOECs for percent germination, germ tube length, and mean sporophyte production (in the Continuous Exposure experiment) were 18 µg/L, 10 µg/L, and 18 µg/L, respectively. There was no significant effect of copper on sporophyte production in the Exposure-Recovery treatment, even at the highest concentrations. The EC50 for percent germination was 54.14 µg/L; the IC50 for germ tube length was 100.88 µg/L. Although the NOEC suggests no greater sensitivity between sporophye production in the Continuous Exposure (CE) experiment and for the short-term endpoints (germination and germ-tube length), mean sporophyte production was inhibited at the lowest copper concentration (5.6 µg/L; Fig. 8a). No statistical difference in sporophyte production was detected at this concentration because of high between-replicate variability. However, the EC50 for sporophyte production in the CE treatment was considerably lower than 50% point estimates calculated for the short-term endpoints (Table 13). This indicates that the magnitude of response and relative sensitivity was greater for sporophyte production in the CE treatment.

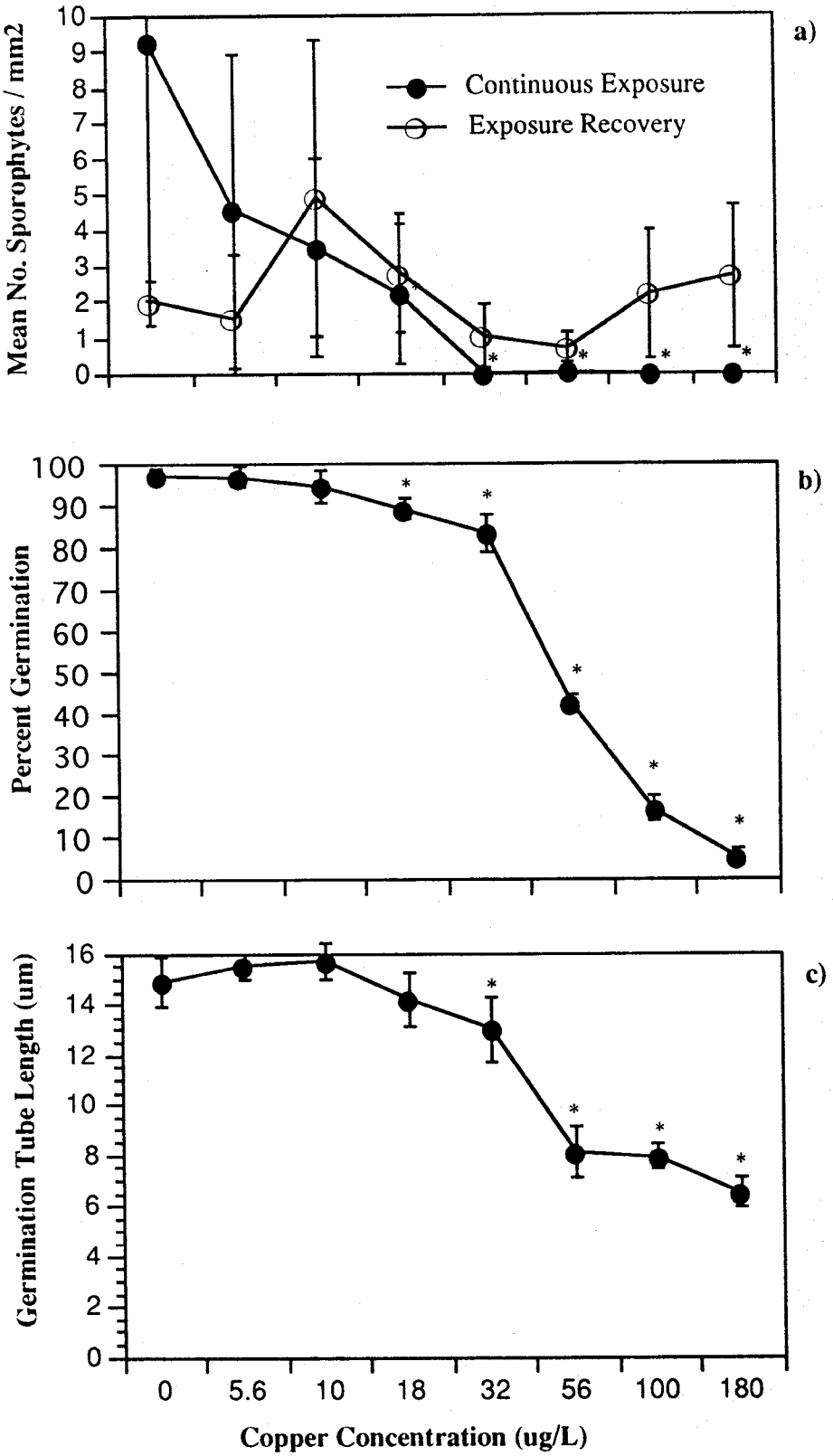
Table 13. Results of 48-hour, Continuous exposure (CE) experiments, and 48-h and 96-h Exposure Recovery (ER) experiments with *Macrocystis pyrifera* using copper chloride (µg/L) as a toxicant. NR indicates no statistical results could be determined.

	Copper Concentrations (µg/L)				
ENDPOINT	**NOEC**	**EC50**	**95% CI**	**IC50**	**95%CI**
Germination (Experiment 1)	10.0	54.14	(51.9-56.5)	-	-
Germ tube length (Experiment 1)	18.0	-	-	110.88	(27.5-137.9)
Sporophyte Production CE (Experiment 1)	18.0	5.5	(2.3-24.3)	-	-
Sporophyte Production 48-h-ER (Experiment 1)	NR	NR	NR	-	-

Figure 8a. Effects of copper chloride on kelp sporophyte production following a 48-h continuous exposure (CE), and following a 48-hr exposure followed by a 17-day recovery period (ER)

Figure 8b. Effects of copper chloride on kelp gametophyte germination following a 48-h exposure.

Figure 8c. Effects of copper chloride on kelp gametophyte growth following a 48-h exposure.


a)
Mean No. Sporophytes / mm2
Continuous Exposure
Exposure Recovery
b)
Percent Germination
c)
Germination Tube Length (um)
Copper Concentration (ug/L)
0
5.6
10
18
32
56
100
180

DISCUSSION

Comparing the relative sensitivity of short-term and longer-term endpoints allows for an estimation of the biological significance of endpoints measured in the 48-h *Macrocystis* germination and growth protocol. These results demonstrate that azide concentrations which inhibit spore germination and gametophyte growth also inhibit kelp reproduction, an ecologically important endpoint. These results are consistent with previous experiments indicating kelp spore germination and initial gametophyte growth may be used as short-term indicators of more chronic effects, although it is clear that vegetative endpoints measured after 48-h exposure are considerably less sensitive than the reproductive endpoint measured after longer exposures. These data do not allow us to determine whether the increased sensitivity in the protocol assessing sporophyte production are due to increased exposure time or a more sensitive endpoint (ie., reproduction); the mechanism of azide toxicity to the various life stages and endpoints used in these experiments are beyond the scope of this study. However, it is possible that the mechanism of azide inhibition of kelp spore germination and germ-tube elongation differs from mechanisms inhibiting kelp reproductive processes. Azide is a cytochrome oxidase inhibitor and therefore blocks ATP production (Kidder and Awayda 1989). Pillai et al. (1992) have shown that germ-tube elongation is a process involving actin microfilaments. Although the precise mechanism is unknown, azide may inhibit actin function through disruption of actin/myosin interactions requiring ATP (Lehninger 1975). Kelp reproduction involves a series of complex biochemical processes requiring ATP, although it is unclear which specific process is affected by azide. There is evidence that many brown alga species rely on a pheromone system to attract sperm to eggs; this may also be disrupted by azide.

Results of the copper experiments were inconclusive primarily due to high between-replicate variability in sporophyte production in the long-term exposures and because of unexplained reductions in sporophyte production in the Exposure Recovery

treatments relative to the the Continuous Exposure treatments. High variability reduced the statistical resolution resulting in a higher NOEC for sporophyte production in the CE treatment. The EC50 for sporophyte production was considerably lower than the point estimates calculated for the germination and germ-tube growth (Table 13). In previous comparisons, we have found sporophyte production in longer-term exposures to be considerably more sensitive to copper (Anderson et al. 1990a). Other researchers have found that reproductive endpoints are generally more sensitive indicators of copper toxicity than vegetative endpoints (see Thursby et al 1993 for review).

There was some recovery in sporophyte production after the intial 48-h copper exposure. It is not clear why sporophyte production in the ER treatment control containers was lower than in the CE treatment controls. It is possible that gametophytes were dislodged when the slides were transferred from the copper exposure chambers after the initial 48-h exposure. Similar transfer proceduress in the second azide test also resulted in lower control sporophyte production in the ER treatment relative to the CE treatment. In this case the reduction in sporophytes was lower, although all experiments demonstrated high between-replicate variability in sporophyte production. Because of the reduction in control sporophyte production in the copper ER control treatment, it can not be concluded that these cultures completely recovered after the initial copper exposure.

Although preliminary, these data suggest that the initial exposure duration determines whether kelp gametophytes exposed to sodium azide are competent to recover and successfully reproduce when transferred to uncontaminated media. Cultures exposed to azide and copper for 48-h were able to recover, while those exposed to azide for 96-h were not. The initial exposure concentration apparently affected the ability of the cultures to recover. For example, there was a reduction in sporophyte production in 48-h-ER cultures exposed to the highest azide concentrations (100 and 180 mg/L). Studies with additional toxicants are necessary to determine whether these results are consistent for compounds having different modes of action.

REFERENCES

Anderson, B. S., Hunt, J. W. (1988). Bioassay methods for evaluating the toxicity of heavy metals, biocides, and sewage effluent using microscopic stages of giant kelp *Macrocystis pyrifera* (Agardh): A Preliminary Report. *Mar. Environ. Res.* 26, 113.

Anderson, B. S., Hunt, J. W., Turpen, S. L., Coulon, A. R., and M. Martin. (1990a). Copper toxicity to microscopic stages of giant kelp *Macrocystis pyrifera*: interpopulation comparisons and temporal variability. *Mar. Ecol. Prog. Ser.* 68, 147.

Anderson, B. S., Hunt, J. W., Turpen, S. L., Coulon, A. R., Martin, M. and F. H. Palmer. (1990b). Marine Bioassay Project Fifth Report: Protocol development and interlaboratory tests with complex effluents. Final Report No. 90-13-WQ. California State Water Resources Control Board, Sacramento, CA, 183pp.

Anderson, B.S., Hunt, J.W., and W.J. Piekarski. In Press (1996). Recent advances in toxicity testing methods using kelp gametophytes. In: Wells, P.G., Lee, K., and C. Blaise (eds.) Microscale Aquatic Toxicology- Advances, Techniques and Practice. Lewis Publishers, Boca Raton, FL.

Chung, I. K., Brinkhuis, B. H. (1986). Copper effects in early life stages of kelp, *Laminaria saccharina. Mar. Poll. Bull.* 17, 213.

Hamilton, M. A., Russo, R. C., and R. V. Thurston. (1978). Trimmed Spearman-Karber method for estimating lethal concentrations in toxicity bioassays. *Environ. Sci. and Tech.*, 11, 714.

James, D. F., Manley, S. L., Carter, M. C., North, W. J. (1987). Effects of PCB's and hydrazine on life processes in microscopic stages of selected brown seaweeds. *Hydrobiologia* 151/152, 411.

Kidder, G. W. III, Awayda, M. S. (1989). Effects of azide on gastric mucose. *Biochem. Biophys. Acta.* 973, 59.

Lehninger, A. L. (1975). *Biochemistry.* (Second edition). Worth, New York. 1104 pp.

Norberg-King, T. (1993). A linear interpolation method for sublethal toxicity: The inhibition cocentration (ICp) approach. National Effluent Toxicity Assessment Center Technical Report 03-93, U.S. EPA Environmental Research Laboratory, Duluth, MN.

Pillai, M. C., Baldwin, J., and G. N. Cherr. (1992). Early development in an algal gametophye: regulation of germination and nuclear events by cytoskeletal elements. *Protoplasma* 170, 34.

Provasoli, L. (1968). Media and prospects for the cultivation of marine algae. In: A. Watanabe and A. Hattori (eds.) Cultures and collections of algae. Proceedings of the U.S.-Japan Conference, Hakone. *Japanese Society of Plant Physiology*, 63.

Thompson, R. S., Burrows, E. M. (1984). The toxicity of copper, zinc, and mercury to the brown macroalga *Laminaria saccharina.* In: *Ecotoxicological Testing For The Marine Environment.* G. Persoone, E. Jaspers, and C. Claus (Eds). State Univ. Ghent and Inst. Mar. Scient. Res., Brendene, Belgium. 2, 588 pp.

Thursby, G. B., Anderson, B. S., Walsh, G. E., and Steele, R. L. (1993). A review of the current status of marine algal toxicity testing in the United States. In: *Environmental Toxicology and Risk Assessment*, ASTM STP 1179, Landis, W.G., Hughes, J.S., and Lewis, M.A. (eds). American Society for Testing and Materials, Philadelphia, PA, 362.

Section 5

Development of the MSD Test Acceptability Criterion to Limit Between-Replicate Variability

INTRODUCTION

In toxicity tests developed by the Marine Bioassay Project (MBP) and others (e.g. Weber et al. 1988, Chapman et al. 1995), test organisms are exposed to toxicants in replicated test containers at a number of toxicant concentrations. Organism responses in test containers are recorded, and these data are subjected to statistical analyses to determine concentrations causing statistically significant effects. While development of biological techniques has been the primary focus of the MBP, toxicity test results are affected by the statistical methodology applied to the data, and the evaluation of appropriate statistics has been the focus of much recent discussion (Chapman et al. 1996, Denton and Norberg-King 1996). Some aspects of statistical analysis are discussed here in the context of toxicity test acceptability and quality control.

At present, the No Observed Effect Concentration (NOEC) is used as the statistic that determines whether an effluent is in compliance with toxicity objectives in regulatory permits. The NOEC is calculated as the highest test concentration producing an organism response that is not significantly different from the response observed in test controls. The NOEC is calculated using hypothesis-testing statistics, such as Analysis of Variance (ANOVA), that compare differences between treatments (toxicant concentrations) with differences between replicates within a treatment. If differences between treatments are sufficiently greater than differences between replicates, the null hypothesis of no difference will be rejected. If within-treatment (between-replicate) variability is high, the test has less power to detect significant differences between treatments, and the level of Type II error (β) may be unacceptably high. In such cases, there is an increased probability of false negatives, in which the statistical test would indicate that the effluent was not toxic when in fact it was. Type I error rates (α) are generally specified in order to limit false positives, that is, determinations that an effluent is toxic when in fact it is not. But since neither the Type II error rate (β) nor statistical power ($1 - \beta$) are specified in current test protocols, the level of between-replicate variation must be controlled to avoid unacceptably high rates of false negative test results.

Variable toxicity testing conditions and concomitant differences in between-replicate variability can create inequities when test results are applied in the regulation of effluent discharges. For example, identical effluent samples tested under different laboratory conditions could produce different NOEC values solely because of differences in between-replicate variability among tests. Such a situation could result in different permit compliance outcomes, even if both tests produced identical mean responses.

A number of statistical approaches are being investigated to address this issue. Ideally, a level of effect considered to be biologically significant would be determined, and statistical tests would be designed to detect that effect size with maximum power. For example, bioequivalence

statistical procedures have distinct advantages in that they allow biologically significant effect sizes to be specified and tested for directly, and they provide natural incentives for the regulated community to seek greater precision in compliance testing. This method is discussed in the context of toxicity testing by Erikson and McDonald (1995). However, hypothesis-testing statistical procedures currently used to analyze toxicity test data suffer from the potential to both ignore large differences if variability is high, and to detect biologically insignificant differences if variability is extremely low. While both of these problems can occur, the test acceptability criteria discussed here address only the issue of unacceptably large variability.

While fundamental improvements in toxicity test statistical procedures are being considered (Chapman et al. 1996, Denton and Norberg-King 1996), the problem of excessive between-replicate variability is being addressed on an interim basis through the adoption of an additional test acceptability criterion in each of the toxicity test protocols. The California State Water Resources Control Board (SWRCB), in conjunction with the US EPA, have adopted such criteria in recent versions of seven toxicity test protocols using west coast test species (Chapman et al. 1995, SWRCB 1996). While previous versions of MBP protocols have controlled variability by setting limits on the ANOVA Mean Square Error (MSE, an indicator of the between-replicate variability), the more recent protocols use the Minimum Significant Difference (MSD) as the basis for determining test acceptability in this regard.

The MSD has advantages as a test acceptability statistic because it incorporates not only the level of between-replicate variability (as indicated by the MSE), but also the number of replicates and the alpha level (Type I error rate). This makes it possible for laboratories to manipulate test design (e.g. increase the number of replicates), if necessary, to meet the MSD test acceptability requirement.

In toxicity tests, the MSD represents the smallest difference between the control mean and a treatment mean (the effect size) that leads to the statistical rejection of the null hypothesis (H_0: no difference). Any effect size equal to or larger than the MSD would result in a finding of statistically significant difference. For example, if the control mean for mysid growth were 80 µg/mysid and the MSD were 20, any treatment with mean mysid weight less than or equal to 60 µg would be significantly different from the control and considered toxic. The same would be true if the MSD were given as a Percent MSD (%MSD) of 25% (or $20 \div 80 = 25\%$). The %MSD for quantal data, such as percent survival or percent normal larval development, is calculated on the basis of arcsine square root transformed data, as indicated in the Data Analysis sections of the test protocols (Chapman et al. 1995, SWRCB 1996). The %MSD statistic (rather than the MSD) was selected for use in test acceptability criteria for the seven west coast protocols to allow more comparability among protocols and to make interpretation of the criteria more intuitive.

The MSD for the commonly used ANOVA with Dunnett's multiple comparison test is calculated as shown below. A similar equation can be derived for MSD values calculated for other statistical tests, such as t-tests.

$$MSD = d * S_W * [(1/n_c + 1/n_t)]^{1/2}$$

where:

d is the critical value for the one-tailed Dunnett's test for the appropriate alpha level, degrees of freedom and Dunnett's "p" value,

S_W is the square root of the within mean square error (an indication of between-replicate variability), and

n_c and n_t are the numbers of replicates for the control and treatments, respectively, assuming equal replication among treatments.

It can be seen from the above equation that the MSD increases with increased between-replicate variability and decreases with increased replication. Therefore, test acceptability can be achieved, if necessary, by decreasing variability or increasing replication. The MSD has been recommended as a means of evaluating toxicity test performance in studies by Thursby (1991), Thursby and Schlekat (1993), and Denton and Norberg-King (1996), and has been recommended as an interim criterion by the Marine Bioassay Project's Scientific Review Committee.

The methods used to determine numerical %MSD test acceptability criteria for toxicity tests using west coast species are described in the following sections. The applications and limitations of these criteria are also considered.

Methods

The %MSD test acceptability criterion for each protocol was based on an evaluation of existing data, and was calculated from the %MSD value attainable by laboratories 90% of the time. Data collection and analysis for the four most widely used tests were conducted by Gwen Starrett of the SWRCB, Debra Denton of the US EPA Region 9, and Robert Smith of EcoAnalysis, Inc., who have presented these results previously (Starrett et al. 1993). The resulting MSD test acceptability criteria are part of a continuing effort to improve toxicity test statistical methodology (Denton and Norberg-King 1996). Further steps are currently underway to evaluate test power, and to evaluate alternative statistical approaches that could be used to

determine compliance with toxicity objectives (Chapman et al. 1996, Erikson and McDonald 1995).

This approach to controlling between-replicate variability is based on determining the %MSD value that was exceeded in only 10% of the tests conducted to date (the 90th percentile %MSD). If this criterion had been in effect while the tests that make up the database were conducted, 10% of those tests would have failed the test acceptability requirement. It is assumed, however, that increasing familiarity with the tests will tend to limit between-replicate variability and reduce the number of test failures. Laboratories that continue to have difficulty meeting this criterion have the option of reducing test %MSDs by increasing the number of replicates in controls and/or treatments, improving control performance, or reducing between-replicate variation where possible through improvements in experimental technique. These options were discussed in detail by Starrett et al (1993), and are described briefly in the following sections.

Database

Data used to generate %MSDs were from reference toxicant tests conducted at a number of laboratories. Data sets for abalone, kelp and sea urchin fertilization tests included more than 100 tests for each protocol. Because the number of tests per laboratory varied, results from each laboratory were randomly subsampled to check for bias (Starrett et al. 1993). There were fewer tests available for the mysid and topsmelt growth and survival protocols, and for the bivalve and echinoderm larval development protocols (Table 13), because these tests have not been widely used in permit monitoring in California. All mysid and topsmelt data were generated at the Marine Pollution Studies Laboratory (MPSL) at Granite Canyon, and because of the limited amount of available data, both reference toxicant and effluent test data were included in the derivation of MSD values for these protocols. All data were obtained from written laboratory data reports and from data submitted electronically to the SWRCB or Regional Water Quality Control Boards. All data were entered into a toxicity information system program (TOXIS) for database management and statistical derivation of MSD values.

All MSD values were derived from Dunnett's Test, and all MSD values were divided by control means to generate %MSD values, which were used in all further computations. Percent MSD values were used to allow more comparability among protocols and to make interpretation of the criteria more intuitive. However, there are statistical limitations associated with dividing the MSD by the control value, which make the %MSD less appropriate than the MSD for some other applications (Smith, draft manuscript).

Table 13. Description of Reference Toxicant Database for Calculating MSD Values.

Toxicity Test Protocol	Reference Toxicant	Number of Laboratories	Number of Tests per Lab	Total Number of Tests
Echinoderm Fertilization	Copper	7	7 - 18	166
Kelp Germination	Copper	8	13 - 24	159
Kelp Germ Tube Length	Copper	8	13 - 24	159
Abalone Larval Development	Zinc	6	10 - 29	135
Bivalve Larval Development	Copper	4	6 - 12	34
Mysid Survival	Zinc/Effluent	1	22	22
Mysid Growth	Zinc/Effluent	1	6	6
Echinoderm Larval Development	Copper	2	6 - 15	21
Topsmelt Larval Survival	Cu/Effluent	1	16	16
Topsmelt Larval Growth	Cu/Effluent	1	10	10

Derivation of 90th Percentile %MSD Values and Acceptability Criteria

The 90th percentile of the %MSD distribution for each protocol was determined graphically from cumulative frequency plots. In cases where laboratory specific bias was detected, plots of the distribution of five random subsets of the data were also presented. Once the 90th percentile value was known, the %MSD test acceptability criterion was determined. For protocols with data from more than 100 tests, the 90th percentile value was simply rounded up to set the criterion value. Rounding was intended to allow a slight margin to improve the likelihood of meeting the test acceptability criterion. For protocols with data from less than 100 tests, the 90th percentile value was used as a basis for the %MSD criterion, but adjustments were made to account for uncertainties or patterns in the smaller data sets. These adjustments are described for each protocol in the following section. As additional data is acquired through wider use of the protocols, these criteria values may be revised as part of continuing efforts to standardize test power.

Results

Analysis of random subsets of the echinoderm fertilization data detected a laboratory-specific bias, so the data distribution of five random subsets was plotted along with the distribution of the actual MSD values (Starrett et al. 1993). The 90th percentile of the actual %MSD distribution (i.e. the point where the cumulative frequency curve crosses the 90% line)

occurs at a %MSD value of 23%, indicating that only 10% of the tests had higher %MSD values (Figure 9). The random subsets of the same data produced a similar curve, with a 90th percentile %MSD value of approximately 25%. A relatively large number of tests were included in this distribution (n = 166), and the %MSD test acceptability criterion was rounded to 25% (Table 14).

The kelp protocol has two test endpoints, spore germination and growth of the germination tube of the developing gametophyte. Both of these endpoints produced similar %MSD values in the 159 tests evaluated. The 90th percentile values were 19% and 17%, respectively (Figure 10). The test acceptability criterion was set at 20% for both of these endpoints (Table 14).

The abalone larval development test produced a 90th percentile %MSD value of 18%, similar to that of the two kelp endpoints (Figure 11). The test acceptability criterion was set at 20% (Table 14). The test acceptability criterion for the bivalve larval development test was 25%, based on a 90th percentile value of 21% (Table 14). Analysis of this data was conducted by Chapman et al (1995), and data were unavailable for graphical presentation for this report.

Mysid survival %MSD data were available from only 22 tests conducted at one laboratory (MPSL). Inspection of the data distribution indicated that the three tests with %MSD values above the 90th percentile were distinct from the rest of the data set (Figure 12a). Including these three tests in the calculation of the 90th percentile %MSD resulted in a value of 46% (Table 14). The test acceptability criterion was set at a %MSD value of 40%, the highest %MSD value that fell below the 90th percentile. This criterion, and others derived from small data sets, should be re-evaluated when more data become available.

The mysid growth (weight) %MSD data were obtained from the smallest data set (n = 6). While mysid survival data were available from 22 tests, many of these tests measured mysid growth in length, rather than weight. Length %MSD values were much lower than those obtained from weight data (Table 6), and were much lower than those obtained from other protocols. For reasons discussed in Section 2, growth in weight is specified as the test endpoint. While the 90th percentile %MSD value for mysid weight is 25% (Figure 12b), the MSD criterion was set at 50% to provide a large margin of error for this endpoint, because weight had been measured in very few tests (Table 14). (Note: The MSD test acceptability criterion for mysid weight was mistakenly given in the EPA and SWRCB procedures manuals as 50µg, which equates to a %MSD of about 68%. The correct value should be 50%, rather than 50µg. However, given the wide margin of error, neither value is likely to be exceeded.) It is clear that more data are required to refine this criterion value. It is anticipated that such a re-evaluation could occur when results from compliance testing with this protocol become available.

Topsmelt survival %MSD data produced a 90th percentile value of approximately 28% (Figure 13a). The topsmelt test with the highest %MSD value in the database (36%) met all

other test acceptability criteria, but it was the first test of this protocol conducted at MPSL. The high level of variability from this test may have been the result of a lack of familiarity with the test organism. More recent experience indicates that %MSD values of 25% should be attainable in at least 90% of tests conducted by qualified laboratories (Table 14).

Topsmelt growth data were available from only 10 tests; therefore only one test had a higher %MSD value than the 90th percentile (Figure 13b). The 90th percentile value was 48%, while all tests had an MSD _ 50%. The %MSD test acceptability criterion was set at 50% (Table 14).

The echinoderm larval development test produced a 90th percentile value of 24% (Figure 14). The %MSD test acceptability criterion was set at 25% (Table 14).

Table 14. Percent MSD values at the 90th percentile of cumulative frequency distributions, and selected %MSD test acceptability criteria for the toxicity test protocols included in the SWRCB and EPA West Coast procedures manuals.

Toxicity Test Protocol	**Total Number of Tests**	**MSD (%) 90th Percentile**	**MSD (%) Test Acceptability Criterion**
Echinoderm Fertilization	166	23	25
Kelp Germination	159	19	20
Kelp Germ Tube Length	159	17	20
Abalone Larval Development	135	18	20
Bivalve Larval Development	34	21	25
Mysid Survival	22	46	40
Mysid Growth (Weight)	6	25	50 *
Echinoderm Larval Development	21	24	25
Topsmelt Larval Survival	16	28	25
Topsmelt Larval Growth	10	48	50

* The %MSD test acceptability criterion for mysid growth should be in units of %, as above, rather than in units of µg, as given in the procedures manuals.

Figure 9. Cumulative Frequency Distribution of %MSD for Echinoderm Fertilization (from Starrett et al. 1993).

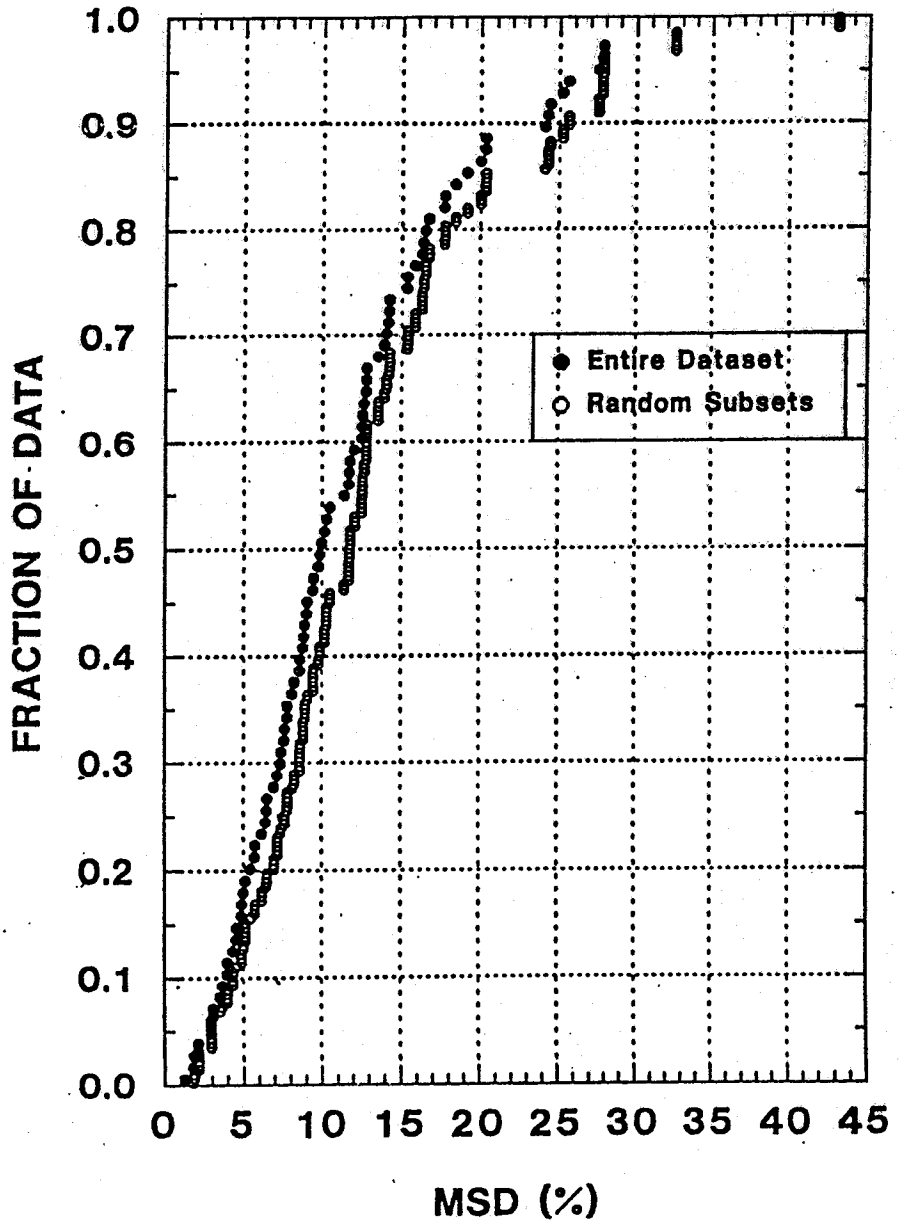


Figure 10. Cumulative Frequency Distribution of %MSD for Kelp Germination and Germ Tube Length (from Starrett et al. 1993).

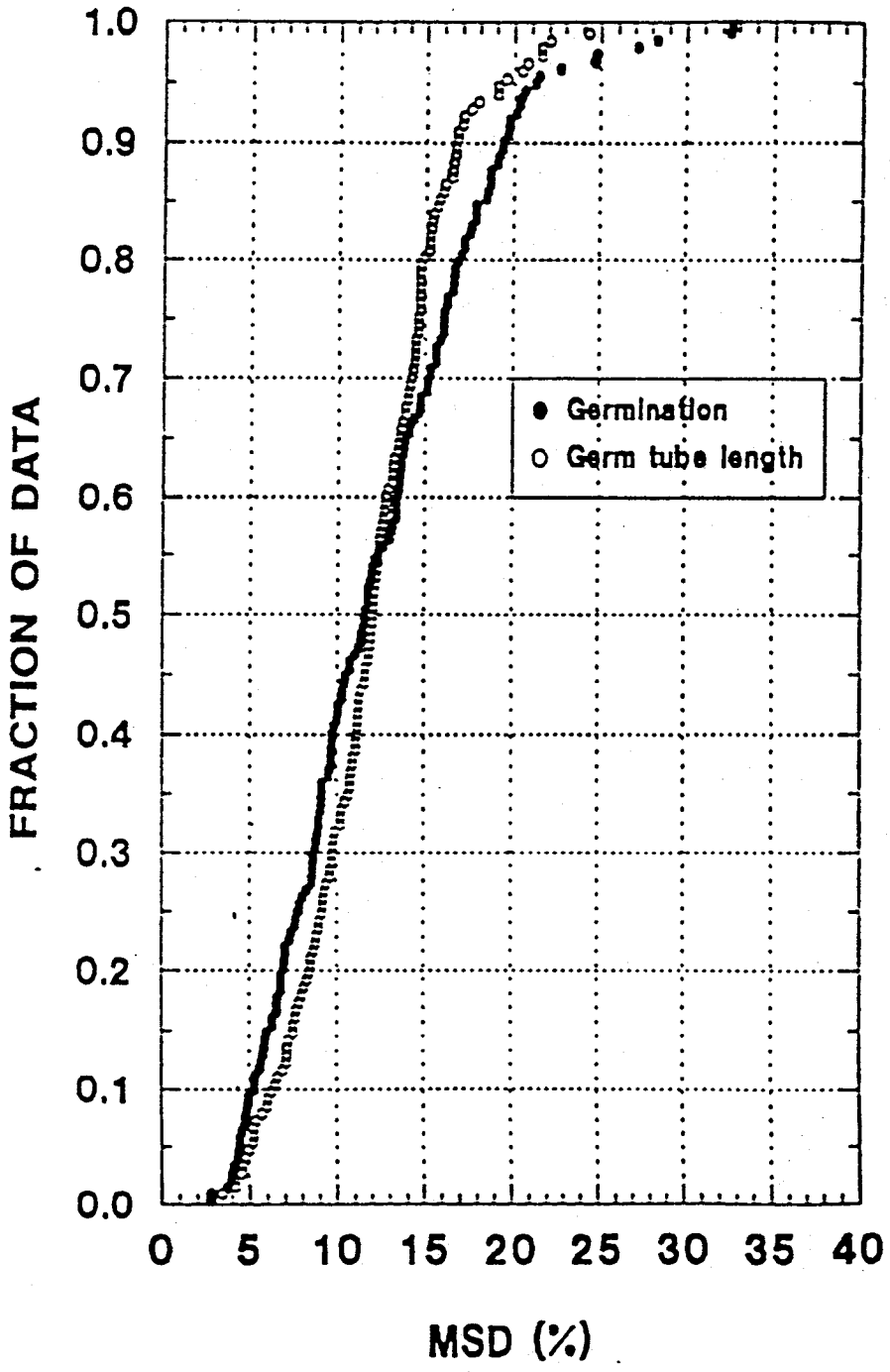


Figure 11. Cumulative Frequency Distribution of %MSD for Abalone Larval Development (from Starrett et al. 1993).

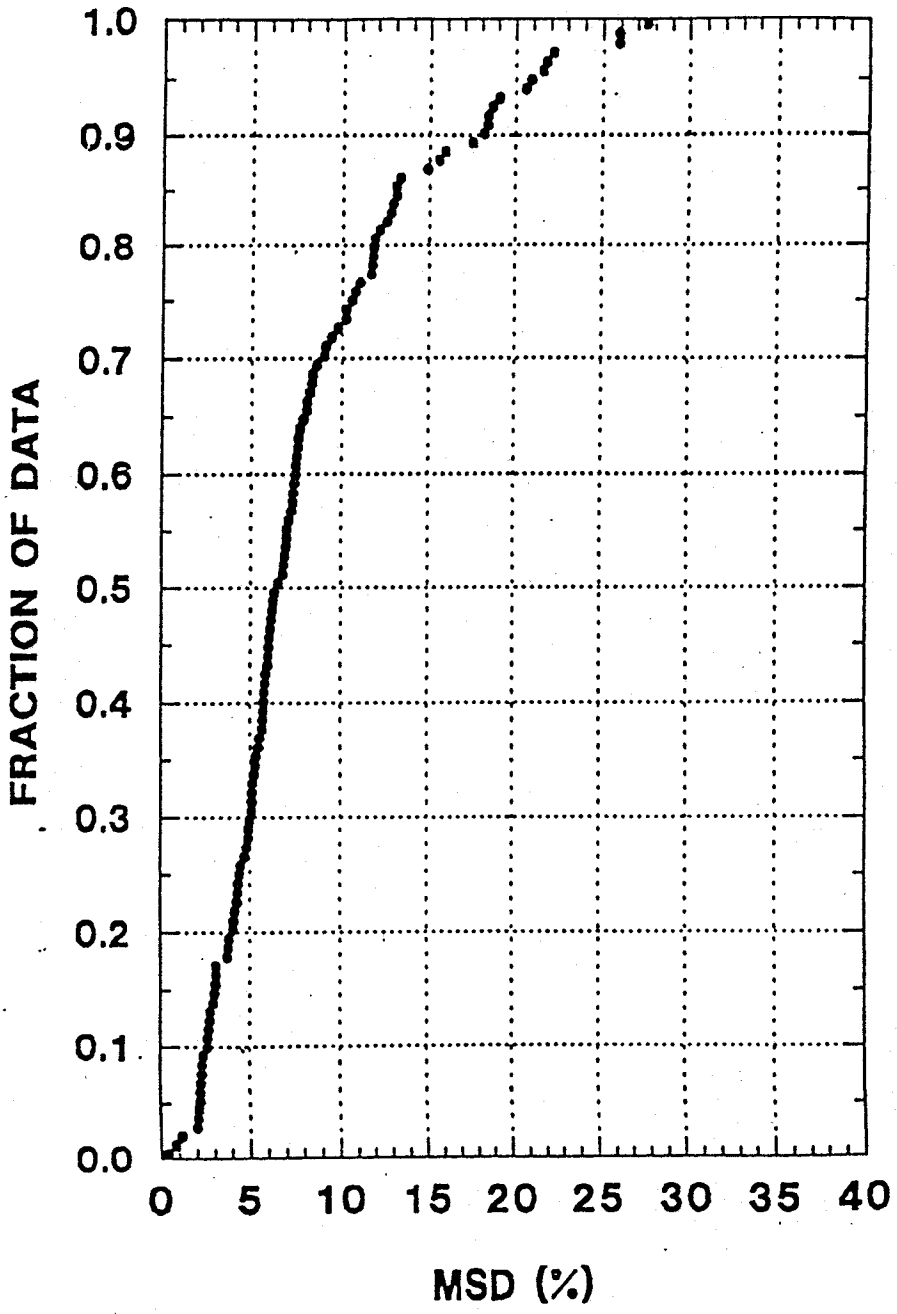


Figure 12a. Cumulative Frequency Distribution of % MSD for Mysid Survival.

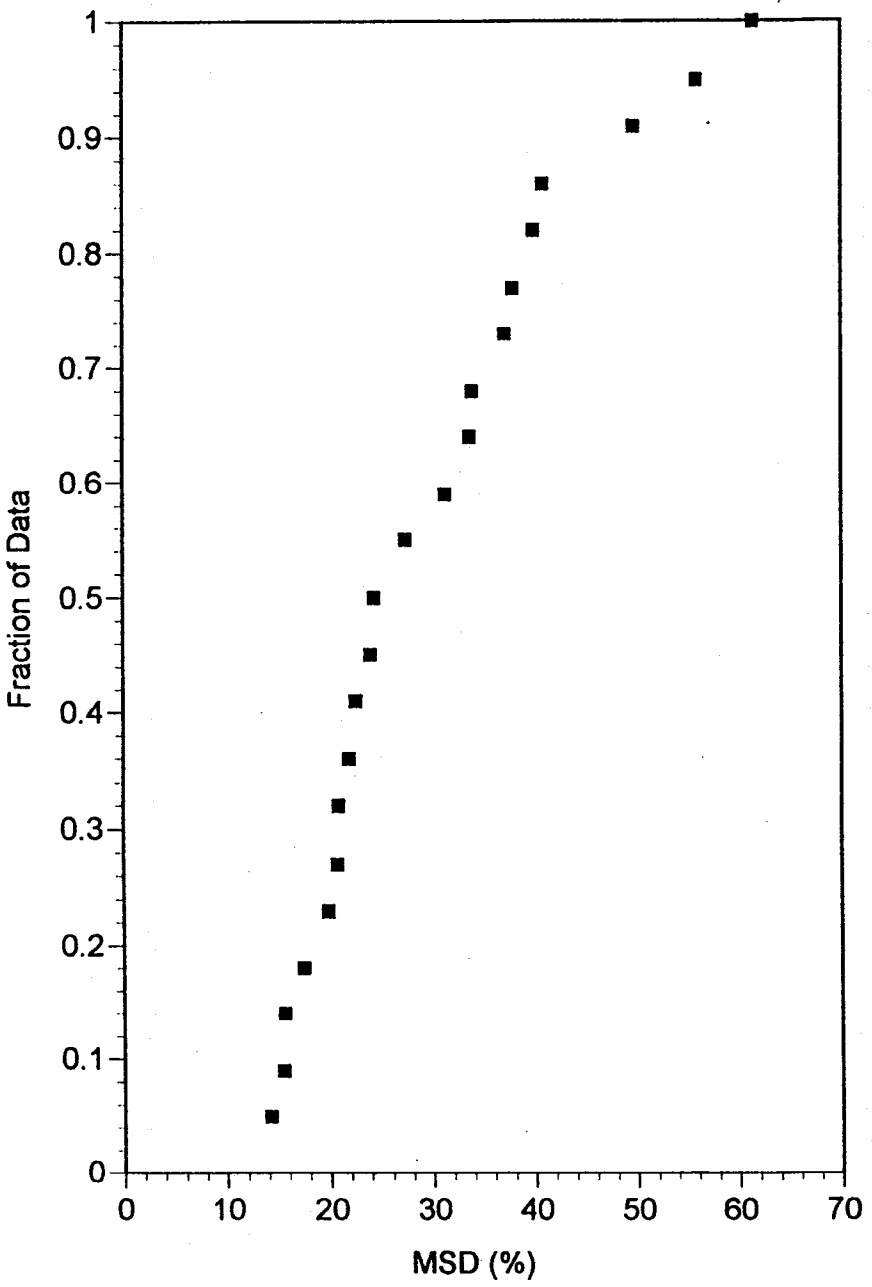


Figure 12b. Cumulative Frequency Distribution of % MSD for Mysid Weight.

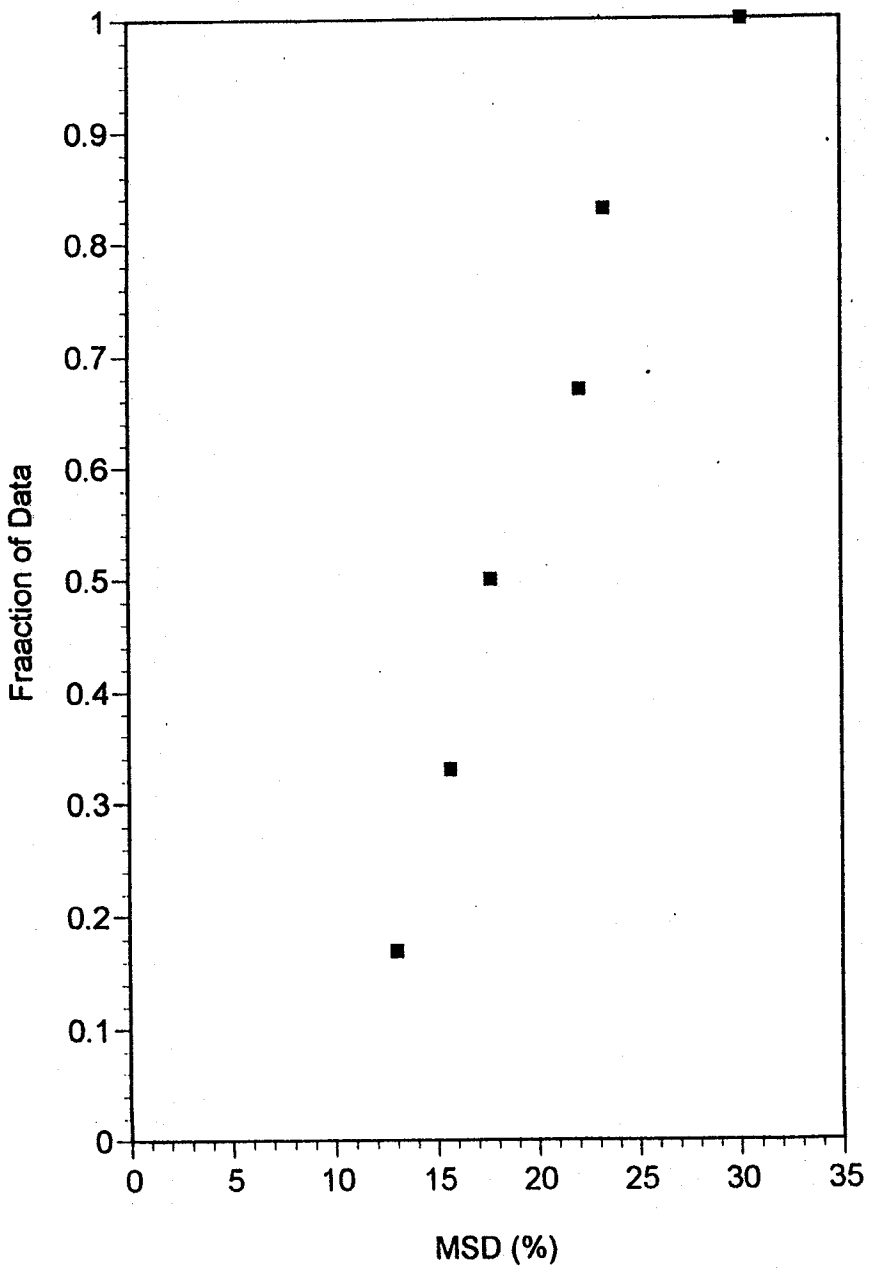


Figure 13a. Cumulative Frequency Distribution of % MSD for Topsmelt Survival.

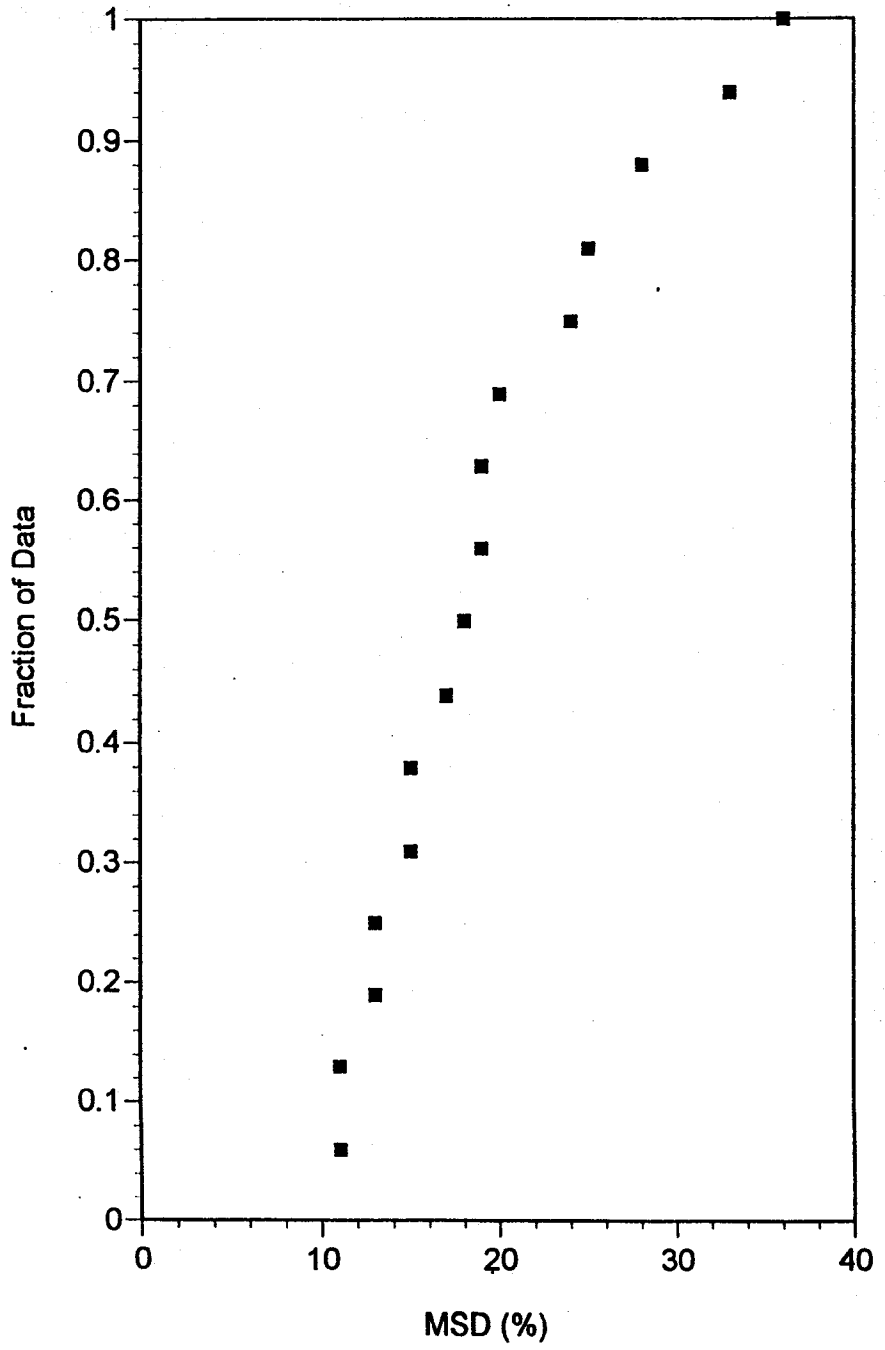


Figure 13b. Cumulative Frequency Distribution of % MSD for Topsmelt Weight.

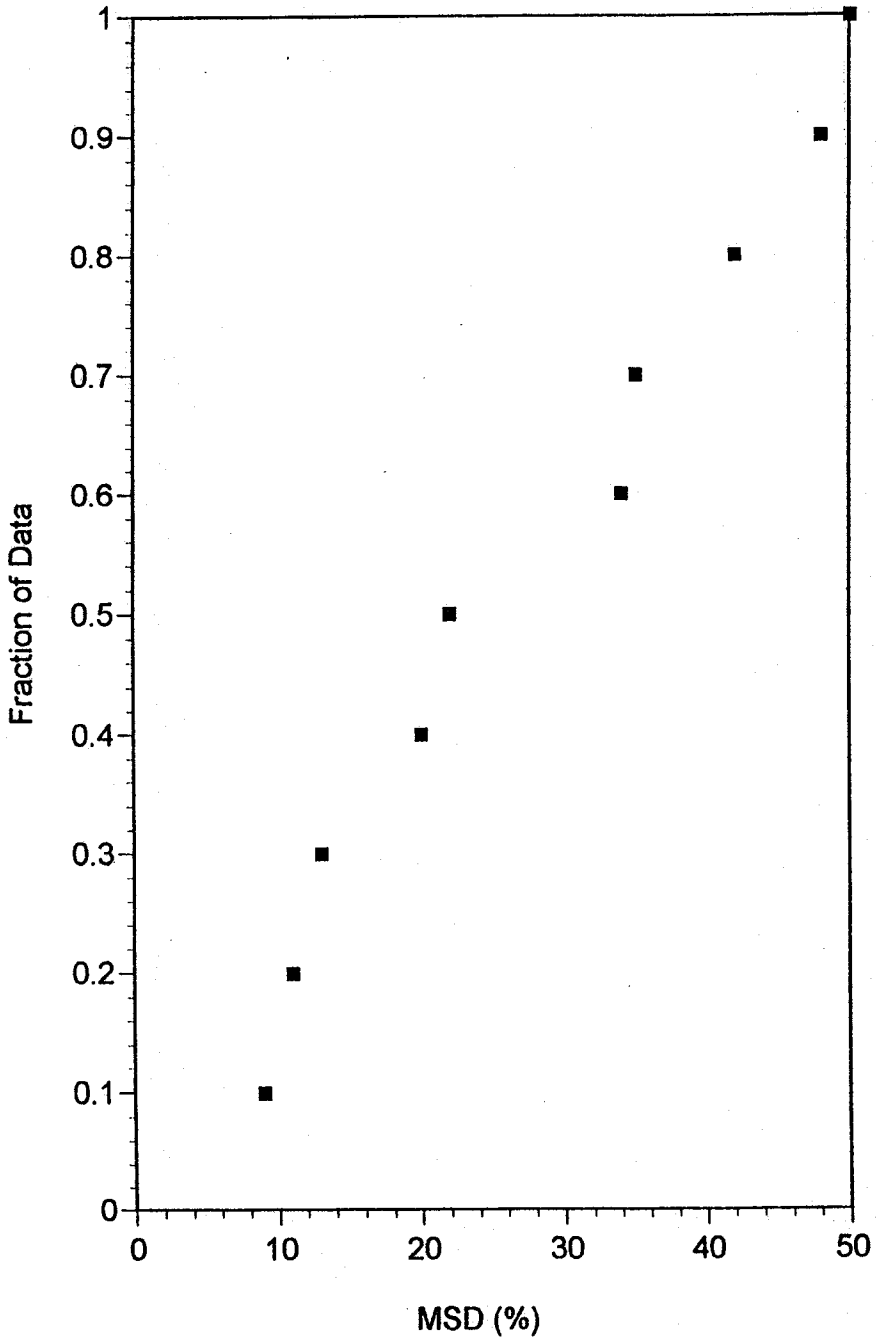
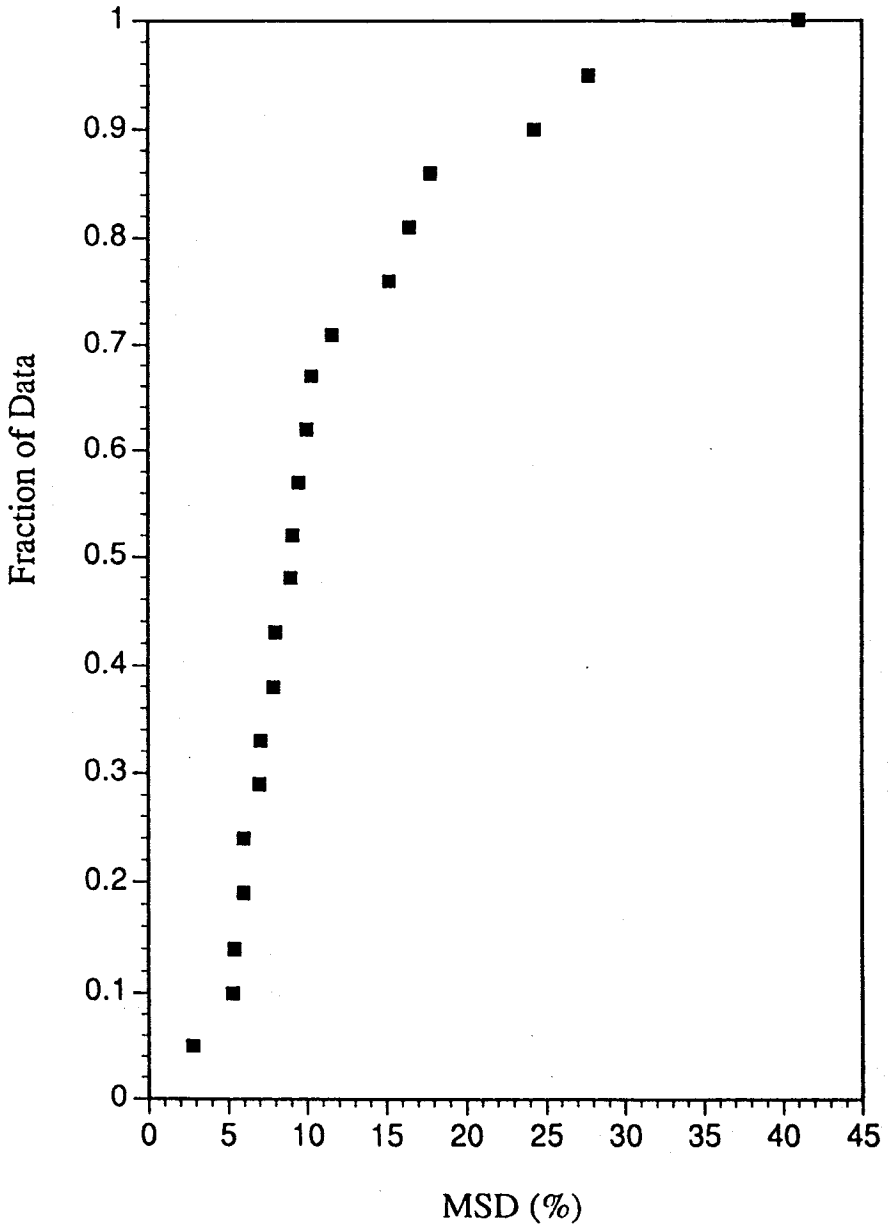


Figure 14. Cumulative Frequency Distribution of %MSD for Echinoderm Larval Development.



Discussion

Statistical methods currently specified in SWRCB (1996) and EPA (1995) toxicity testing procedures manuals have a decreased ability to detect toxicity when variability among laboratory replicates is high. This may create a situation where there is less incentive for the regulated community to produce precise toxicity test data, since imprecise data is more likely to lead to compliance with toxicity objectives specified in discharge permits. To limit the extent to which variable data affects both test results and subsequent permit compliance, regulators have used the %MSD (percent Minimum Significant Difference) to control allowable levels of between-replicate variability (Chapman et al. 1995, SWRCB 1996). Participants in this effort agree that this is a necessary interim step while seeking consensus support for alternative statistical techniques that will provide incentives for data quality.

The %MSD test acceptability criteria included in recent toxicity test protocols (Chapman et al. 1995, SWRCB 1996) have been based on the graphical analyses described here. The databases for three of the test protocols (abalone, kelp, and echinoderm fertilization) are sufficiently large to estimate the level of precision that qualified laboratories are capable of obtaining under routine testing conditions. There are relatively few data available to predict the rate of test acceptability that can be obtained when conducting the other four protocols specified in the recent manuals. Close observation by regulators and feedback from testing laboratories will be necessary to determine when and how %MSD test acceptability criteria for the mysid, topsmelt, bivalve and echinoderm larval development tests should be re-evaluated. Clearly, more data is necessary to adequately define %MSD criteria for mysid and topsmelt growth endpoints. This data would be most appropriate if it were produced by a number of laboratories conducting tests for permit compliance purposes.

While the %MSD criteria for some protocols may require revision in the near future, effort may be more effectively spent by pursuing alternative approaches for regulating test power. The current %MSD criteria, while useful and necessary, are an interim solution. The use of test acceptability criteria in general has statistical ramifications if test results are eliminated because of random events rather than procedural errors. The elimination of data that vary randomly beyond criteria limits may alter the underlying Type I and Type II error levels of the statistical tests applied to the data. Such cases would require special techniques to adjust for biases of the means and variances (Smith, draft document; El-Shaarawi et al. 1989). In addition, the %MSD test acceptability criteria, while limiting the occurrence of false negatives, do not prevent small differences of uncertain biological significance from being considered statistically significant in cases where between-replicate variability is small.

Efforts are currently underway to evaluate test power for a variety of protocols under a variety of test conditions. Achieving a standard level of power for each protocol would be a more direct method of regulating test precision. Power curves have been constructed to show the relationship between the MSD, the amount of change between the control and treatment, and statistical power. With this information, a regulator could select a desired level of change from the control, for example 20%, and could read from the power curve the MSD value required to detect that change with a given level of statistical power (Denton and Norberg-King 1996; Smith, draft manuscript). While statistically sound, this method still leaves the regulator in the position of having to enforce a certain level of test precision.

Alternatively, bioequivalency tests, currently under consideration in a number of forums (McDonald and Erikson 1994, Chapman et al. 1996), may provide the best means of achieving optimal test power by providing incentives for both regulators and the regulated community to obtain the most precise data possible. Presently mandated statistical tests seek to disprove the null hypothesis that there is no difference between a treatment and the control. If the difference between treatment and control means is large relative to variation among replicates, the null hypothesis is rejected and there is a determination of statistical difference. With the bioequivalence test approach, the null hypothesis is that there is no difference between the treatment and a designated deleterious effect level (which might be, for example, 80% of the control value). Compliance with the toxicity objective would be achieved if the test established that the null hypothesis was false, indicating that the treatment mean was significantly less toxic than the established deleterious effect level. To demonstrate this difference, there would be an incentive to characterize the treatment mean and the control mean (of which the effect level is a fraction) as precisely as possible, so that confidence intervals around the two means would not overlap. For example, if the designated effect level were 80% of the control value, an effluent with a mean of 85% of the control value would meet its compliance objective if there were very low variability, such that the 85% treatment mean were shown to be statistically different from the 80% designated deleterious effect level. While some effort would be necessary to establish a deleterious effect level, it must be recognized that such a level is now implicitly established through currently applied statistical methods, and that this level is different for every test, depending on between-replicate variability.

Discussion of these and other alternative methods will continue. While the present report has described necessary interim measures to control test precision, it may be beneficial for all involved to focus future efforts on fundamental changes to statistical methods, rather than pursuing incremental improvements in test acceptability criteria regulating test performance.

References

Chapman, G. A., D. L. Denton, and J. M. Lazorchak. 1995. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to west coast marine and estuarine organisms. First Edition. EPA/600/R-95-136. US Environmental Protection Agency, Cincinnati, Ohio.

Chapman, G.A, B. Anderson, A. Bailer, R. Baird, R. Berger, D. Burton, D. Denton, W. Goodfellow, M. Heber, L. McDonald, T. Norberg-King, P. Ruffier. 1996. Methods and appropriate endpoints: Discussion synopsis. In: Grothe, D.R., K.L. Dickson, and D.K. Reed-Judkins (Eds.), Whole effluent toxicity testing: an evaluation of methods and prediction of receiving water impacts. SETAC Special Publication, SETAC Foundation for Environmental Education, Pensacola, FL.

Denton, D.L. and T.J. Norberg-King. 1996. Whole effluent toxicity statistics: a regulatory perspective. In: Grothe, D.R., K.L. Dickson, and D.K. Reed-Judkins (Eds.), Whole effluent toxicity testing: an evaluation of methods and prediction of receiving water impacts. SETAC Special Publication, SETAC Foundation for Environmental Education, Pensacola, FL.

El-Shaarawi, A.H., P.B. Kauss, M.K. Kirby, and M. Walsh. 1989. Inferences about the variability of means from censored data. Environ. Monitoring and Assessment 12: 295-304.

Erickson, W.P. and L.L. McDonald. 1995. Tests for bioequivalence of control media and test media in studies of toxicity. Environ. Toxicol. and Chem. 14: 1247-1256.

McDonald, L.L. and W.P. Erickson. 1994. A shift in paradigm for statistical analysis in risk assessment. Abstract, 15th Annual Meeting, Society of Environmental Toxicology and Chemistry.

Schimmel, S.C., B.D. Melzian, D.E. Campbell, C.J. Strobel, S.J. Benyi, J.S. Rosen, H.W. Buffum, and N.I. Rubinstein. 1994. Statistical Summary EMAP-Estuaries Virginian Province - 1991. EPA /620/R-94/005.

Smith, R.W. Draft manuscript on hypothesis testing statistical techniques. In: Statistical issues in whole effluent toxicity testing: problems and proposed solutions. Draft Report to the Sothern California Toxicity Assessment Group.

Starrett, G.L, D.L. Denton, and R.W. Smith. 1993. Sensitivity of toxicity test protocols and the need for additional quality assurance requirements. Abstract, 14th Annual Meeting, Society of Environmental Toxicology and Chemistry.

SWRCB. 1996. Procedures manual for conducting toxicity tests developed by the Marine Bioassay Project. California State Water Resources Control Board 96-1WQ, Sacramento, CA. 639 pp.

Thursby, G.B. 1991. Complex effluent toxicity test performance: control response and minimal detectable differences from a single laboratory. EPA Contract No. 68-C8-0066, WA No. C-2-41 (O). Report submitted to US EPA Office of Water Enforcement and Permits, Washington D.C.

Thursby, G.B. and C.E Schlekat. 1993. Statistical analysis of 10-day solid phase toxicity data for amphipods. Abstract, 14th Annual Meeting, Society of Environmental Toxicology and Chemistry.

Weber, C.I., W.B. Horning, D.J. Klemm, T.W. Neiheisel, P.A. Lewis, E.L. Robinson, J. Menkedick, and F. Kessler. 1988. Short-term methods for estimating chronic toxicity of effluents and receiving waters to marine and estuarine organisms. EPA/600/4-87/028. U.S. Environmental Protection Agency, Cincinnati, OH.

SECTION 6

TECHNICAL SUPPORT AND ORGANISM SUPPLY

Topsmelt Supply

One of the primary goals of the MBP is to promote use of the the various toxicity test protocols by establishing a network of reliable test organism suppliers. Supply has been established for three of the four MBP protocols. Abalone (*Haliotis rufescens*) supply has been accommodated by the various commercial operations who culture this species for the seafood industry. Kelp (*Macrocystis pyrifera*) reproductive blades and mysids (*Holmesimysis costata*) are available year-round from various suppliers who collect these species from the field. Topsmelt (*Atherinops affinis*) larval supply has been the most problematic because of the technical difficulties involved in year-round laboratory culture of a species which normally only spawns in spring and early summer. We have demonstrated that this species may be spawned year-round in laboratory culture (Anderson et al. 1994) and have recently been providing information to test organism suppliers interested in culturing this species.

In an effort to establish supply of topsmelt larvae for toxicity test purposes, MBP staff met with personnel from Aquatic BioSystems (ABS) of Fort Collins, Colorado in late 1994 to discuss the possibility of establishing a breeding population at their facility. ABS provide a number of test species, including *Menidia*, a genus with culture requirements similar to topsmelt. Protocols and recent publications on topsmelt culture were provided to ABS, and a culture system was assembled at their facility. Twenty-five fish were shipped in December, 1994 and these began spawning within two weeks. A second group of 300 adult fish were transported to Colorado in June 1995 using a tanker truck. These fish arrived safe, but unfortunately all died within three weeks due to a system failure leading to high temperatures and ammonia concentrations. The culture system was modified to safeguard against extreme temperatures and ammonia.

At Granite Canyon, considerable effort was devoted to defining appropriate shipping conditions to ensure safe arrival of adult topsmelt shipped via overnight courier. The primary problem encountered during shipment was oxygen supersaturation. Fish shipped overnight require adequate dissolved oxygen levels, but this needs to be balanced to prevent DO supersaturation which is lethal to fish. After several failed efforts, procedures to ensure adequate aeration were developed. This involved balancing number of adult fish per bag (5 fish per 5 liters seawater), temperature (17°C), and DO (pure O_2 in 1 liter head space at top of each shipping bag). In addition, MBP and ABS staff collaborated on culture methods to ensure good survival in the artificial seawater formula utilized at the ABS facility. Because the fish were originally cultured in natural seawater

at Granite Canyon, it was discovered that a 24-h adjustment period was required to acclimate fish to artificial seawater.

ABS now has a population of 1500 topsmelt which are near reproductive age and size. These are fish which have been cultured from larvae hatched at their facility in February, 1995. It is anticipated the current population of fish at ABS will be spawnable in summer 1996. An additional population of 25 adult topsmelt have recently (February, 1996) been shipped to ABS to supplement their culture population. It is anticipated that more adult fish will be sent as needed in June, 1996.

In December 1995, MBP staff were contacted by Aquatic Indicators (AI) of Saint Augustine, Florida regarding topsmelt supply. AI also supply *Menidia* to testing laboratories. After initial discussions, protocols and recent publications on topsmelt culture were provided to AI, and a culture system was assembled at their facility. In December, 1995 approximately 2000 topsmelt embryos were shipped to AI, and larvae hatched from these are now being cultured at their facility (H. Hickock, Aquatic Indicators, personal communication). We plan to continue shipping adult fish to this group throughout the spring to establish a viable culture for year-round topsmelt supply.

It is anticipated that establishment of these two suppliers should provide a reliable supply network of this species for NPDES monitoring purposes on the west coast.

TOXIS Workshops and Development of Statewide Chronic Toxicity Database

In 1991 we contracted a private company (EcoAnalysis) to develop TOXIS, a toxicity database/statistics package designed for entering, analyzing, and retrieving Whole Effluent Toxicity (WET) test data. Since receiving the first working version of the program, MBP staff used the program on a test basis to help facilitate development of a fully operable program for use statewide. MBP staff continued to work closely with the EcoAnalysis staff to further enhance the efficiency of the program. In addition to initial refinements, State Board staff requested development of a Compliance Module to be added to the database so that compliance data could be rapidly assessed using established Ocean Plan Toxicity Objectives. TOXIS has evolved through continued use and is now used statewide by a number of municipal and industrial dischargers.

EcoAnalysis was contracted to conduct a series of workshops demonstrating the utility of the TOXIS database program for analyzing WET testing data. Workshops were designed to provide information to State and Regional Board staff, particularly for permit writers who routinely review WET test data. State Board staff established a structured format for test data to facilitate data incorporation into a larger statewide database. Five

separate workshops were conducted by EcoAnalysis to instruct staff from the following Regional Water Quality Control Boards:

TOXIS Workshops

Workshop Number	Date	Participants
1	June 7, 1994	San Diego Regional Board (9)
2	June 12, 1994	Los Angeles (4) and Central Coast Regional Boards (3)
3	July 26, 1994	Central Valley Regional Board (5)
4	August 9, 1994	San Francisco Regional Board (2)
5	September 20, 1994	Santa Ana Regional Board (8)

Statistics Issues

In addition to answering technical questions regarding the MBP protocols, project staff have also participated in meetings and workshops held to address statistical issues concerning the use of toxicity test data. In September 1994, project staff attended a statistical workshop organized by the Southern California Toxicity Assessment Group (SCTAG), and presented a summary of current MBP activities. In November, 1994, while at the National SETAC meeting, project staff participated in meetings at EPA's Region 8 offices in Denver, Colorado. Issues related to analyses of toxicity test data were discussed.

In September 1995, project staff also participated in a national conference on WET testing sponsored by the SETAC Foundation for Environmental Education. The conference was convened for a week in Pellston Michigan and was attended by 50 experts in WET test issues representing industry, government and academia. Workshop participants were divided into four technical sessions addressing specific WET issues: (1)

Methods/ Appropriate Endpoints, (2) Effluent Toxicity Test Variability, (3) Field Assessments, (4) Predicting Receiving Water Impacts from Effluent Toxicity. After initial presentations and consideration of issues papers related to the various subjects, the participants separated into the four technical sessions. MBP staff participated in Session 1 on Methods and Appropriate Endpoints. The objective of this session was to evaluate currently used statistical and biological endpoints, the relationship between biological significance (ie. effects) and statistical significance, and issues related to detection and quantitation levels for marine and freshwater tests. The strengths and weaknesses of various statistical tests used in WET compliance monitoring and precision estimation were also discussed. MBP staff provided formal review of two of the technical issues papers and reviewed the chapters from each technical sessions. The proceedings of this workshop will be included in a SETAC book to be published in Summer 1996.

Toxicity Test Protocol Workshops

As part of ongoing statewide training for WET testing, MBP staff participated in an EPA-sponsored workshop on chronic toxicity testing at the EPA's Region IX Richmond Field Station in June 1994. Using a combination of slide presentations and hands-on participation in spawning and endpoint evaluations, MBP staff demonstrated the kelp, abalone, and sea urchin fertilization protocols to EPA and State and Regional Board staff.

EPA Bioremediation Workshop

MBP staff will be participating in an EPA-sponsored workshop on bioremediation metabolite toxicity in March 25-26 1996, in Alexandria Virginia. Bioremediation of toxicants such as petroleum products, PCBs and other metal and organic contaminants is a growing international field in toxicology. Because some of the metabolites produced by microbes are often more toxic than the parent compounds, there is growing concern over the use of this tool in environmental toxicology. The U.S. EPA is considering implementation of effluent discharge limits to regulate discharge of bioremediation products, including chemical and biological toxicity screening. EPA, in conjunction with the Canadian government, is sponsoring a workshop to address various aspects of bioremediation toxicity and pathogenicity. MBP staff were invited to this workshop to discuss appropriate toxicity test species and endpoints for this application. The final product of this workshop will be an issues paper discussing the state of the science, objectives for government regulation, and further research needed.

The Marine Bioassay Project

California State Water Resources Control Board

Francis H. Palmer, D.Env.
Matthew W. Reeve

California Department of Fish and Game

Project Managers:
Mark D. Stephenson
H. Max Puckett

University of California, Santa Cruz

Principal Investigators:
Brian S. Anderson
John W. Hunt

Project Staff:
Witold Peikarski
Bryn M. Phillips

Scientific Review Committee

Steven Bay, SCCWRP
Thomas Dean, Ph.D., Coastal Resources Associates, Inc.
JoEllen Hose, Ph.D., Occidental College