Overview of Freshwater and Marine Toxicity Tests:

A Technical Tool for Ecological Risk Assessment

April 2004

Ecotoxicology Program
Integrated Risk Assessment Branch
Office of Environmental Health Hazard Assessment
California Environmental Protection Agency

Overview of Freshwater and Marine Toxicity Tests

California Environmental Protection Agency
Office of Environmental Health Hazard Assessment
Integrated Risk Assessment Branch
Ecotoxicology Program

April, 2004

This document was prepared under a contract between the California Environmental Protection Agency (Cal/EPA) Office of Environmental Health Hazard Assessment and the Department of Environmental Toxicology, University of California, Davis

Authors

Brian Anderson
Patricia Nicely
Kristine Gilbert
Rosamaria Kosaka
John Hunt
Bryn Phillips

Department of Environmental Toxicology, University of California, Davis

Reviewers

James M. Donald, Chief, Ecotoxicology Unit
Karen Randles, Associate Toxicologist
Barbara Shayne Washburn, Staff Toxicologist

Cal/EPA Office of Environmental Health Hazard Assessment

Preface

This document is one of a series of technical resources and tool prepared by the Ecotoxicology Unit (OEHHA) as part of our goal of advancing the science and practice of ecological risk assessment. It is anticipated that the ecological risk assessment community will utilize this document as a preliminary tool to conveniently review existing scientific information on Freshwater and Marine Toxicity Tests. Efforts have been made to ensure that the information in this document accurately reflects the original source. However, users should refer to the original publication in order to thoroughly understand the tests and any appropriate limitations on their use. Cal/EPA, OEHHA and the University of California Regents are not responsible for damages of any kind resulting from the use of or reliance on this information by risk assessors or risk managers.

TABLE OF CONTENTS

This page is intentionally left blank

INTRODUCTION

Environmental managers responsible for assessing the ecological integrity of aquatic resources in California rely on a number of assessment tools including chemical analysis of water, sediment, and tissue; biological assessments; and toxicity tests. Toxicity tests are an important component for assessing the impact of chemicals on aquatic ecosystems because they indicate toxic effects of complex chemical mixtures. In aquatic toxicity tests, groups of selected organisms are exposed to test materials (water or sediment samples) under defined conditions to determine potential adverse effects. A number of standardized toxicity test protocols have been developed for determining toxicity of chemicals to aquatic species. Detailed guidance manuals for marine and freshwater toxicity tests are available from the United States Environmental Protection Agency (U.S. EPA) and other entities such as the American Society for Testing and Materials (ASTM). These protocols provide guidance on application of toxicity tests for assessing toxicity of single chemicals, complex effluents, and ambient samples of water or sediment.

The following document is intended to provide an overview of the various standardized aquatic toxicity test protocols available for hazard assessment. Methods for evaluating the toxicity of water and sediment samples from marine and freshwater environments are described in Part I and Part II, respectively. Relative sensitivities of the various protocols are discussed in terms of their responses to single chemicals in reference toxicant exposures, and are also compared using studies of ambient water and sediment samples. Methods for assessing bioconcentration and bioaccumulation of chemicals in water column and sediment test matrices are also described. In addition, the strengths and limitations of the various protocols are discussed using examples from the scientific literature, and factors that may influence or confound interpretation of toxicity test results are described. Guidance for applying water column and sediment toxicity tests in environmental assessments is also provided; this guidance emphasizes considerations for selecting the different test protocols for use in Ecological Risk Assessments, but is also applicable for hazardous waste site evaluations, Natural Resource Damage Assessments, and other situations requiring toxicity evaluations.

This document is intended to familiarize environmental managers with one of the tools used by ecotoxicologists for environmental assessments, but is not intended to be a comprehensive review of aquatic toxicity testing methods. Although it is recognized that a variety of other non-standardized toxicity test methods are used in ecotoxicologic research, emphasis is placed on standardized protocols provided by the U.S. EPA and ASTM, because these are the tests most commonly used in regulatory applications. In addition, species and protocols relevant for California waters are emphasized. Part III of this document gives recommendations for using aquatic toxicity tests as part of a weight-of-evidence approach in integrated ecotoxicologic studies.

PART I. MARINE TOXICITY TEST METHODS

Marine Water Column Toxicity Tests

The California State Water Resources Control Board (SWRCB 1996) and the U.S. EPA (1995a) list seven marine toxicity test protocols considered to be appropriate for Whole Effluent Toxicity (WET) tests as part of the National Pollutant Discharge Elimination System toxicity compliance monitoring program in California. As concern for water quality has shifted from point source to non-point pollution sources, some of these protocols have also been used in marine and estuarine ambient toxicity monitoring programs. Four of these protocols were developed as part of the State Water Board's Marine Bioassay Project (*Haliotis rufescens, Holmesimysis costata, Atherinops affinis, Macrocystis pyrifera*); three were developed by other researchers (for *Mytilus galloprovincialis* and *Strongylocentrotus purpuratus*). All of these test protocols were developed using species indigenous to California because of their ecological relevance, and due to concerns over the accidental introduction of non-native species. The following discussion gives brief descriptions of these test procedures and where appropriate, lists alternate species that have also been tested using these protocols.

Marine Embryo-larval Water Toxicity Tests

Short-term Embryo-larval Water Toxicity Tests: Purple Sea Urchin (*Strongylocentrotus purpuratus*), Red Abalone (*Haliotis rufescens*), Bay Mussel (*Mytilus galloprovincialis*), and Alternate Species

Compiled from U.S. EPA 1995a, SWRCB 1996

Haliotis rufescens and *S. purpuratus* are found in marine environments along the Pacific coast. Adults of these species are ecologically important as grazers of marine algae and as food for sea otters and predatory invertebrates. *Mytilus galloprovincialis*, a sessile filter feeder, is found in estuarine and low-wave energy marine environments. It is also an important prey item, and can form large aggregations that provide habitat for

other organisms. *Mytilus galloprovincialis* and *H. rufescens* are also valued by humans as food items, and are harvested commercially and for sport.

Adults of these three species are used as brood stock for embryo-larval toxicity tests. Brood stock can be collected from wild populations, but are typically supplied by commercial labs that can provide reliably reproductive individuals. Embryo-larval tests are initiated by inducing gamete release (spawning) in male and female brood stock, and combining the eggs and sperm to form the embryos used in testing. Spawning is induced differently in these species: *H. rufescens* is spawned in cool, aerated seawater to which hydrogen peroxide and Tris reagent are added; *S. purpuratus* is injected with potassium chloride; and *M. galloprovincialis* is subjected to a warm-water treatment. Once a sufficient number of gametes have been produced, eggs and sperm are combined, fertilization occurs, and embryo densities are determined.

Tests are typically performed in small, covered, glass containers to which test solution (10 to 200 mL) is added. Test solutions can consist of marine samples, salted fresh or estuarine samples, seawater/saltwater controls, and reference toxicant controls. Four to five replicate containers of each test solution are inoculated with a known density of embryos that develop into motile larvae over the duration of the test. The final density of these embryos ranges from 10 to 25 per mL, depending on the species being tested. These static, non-renewal tests are terminated after 48 hours (*H. rufescens*, *M. galloprovincialis*) or 72-96 hours (*S. purpuratus*) by the addition of buffered formalin. The endpoint, percent normal development, is determined by counting normally and abnormally developed larvae using an inverted compound microscope.

Alternative species can be used in place of *M. galloprovincialis* and *S. purpuratus*. Pacific oyster (*Crassostrea gigas*) has been used in place of *Mytilus* in the bivalve test. Sand dollar (*Dendraster excentricus*) has been used in place of purple urchins with no change in the testing procedure.

An echinoid (purple urchin or sand dollar) test with fertilization success as the endpoint is commonly used. In this test, known densities of sperm are added to the test solutions, and known densities of eggs added 20 minutes later. Fertilization is allowed to occur for 20 minutes, after which time the test is terminated and fertilization success is determined by the appearance of the fertilization envelope.

Table 1. Test conditions for conducting a water column toxicity test for larval invertebrate development: *Haliotis rufescens, Strongylocentrotus purpuratus, Dendraster excentricus, Mytilus galloprovincialis,* or *Crassostrea gigas.*
(Compiled from SWRCB 1996)

Parameter	Conditions
Test type	Water column, static, non-renewal
Temperature	15°C: *S. purpuratus*, *D. excentricus*, *H. rufescens* 15 or 18°C: *M. galloprovincialis* 20°C: *C. gigas*
Salinity	30‰: *M. galloprovincialis*, *C. gigas* 34‰: *S. purpuratus*, *D. excentricus*, *H. rufescens*
Light quality	Ambient laboratory light
Illuminance	10 to 20 μE/m^2/s (ambient laboratory levels)
Photoperiod	16 light:8 dark
Test chamber	30-mL glass vial: *S purpuratus*, *D. excentricus*, *M. galloprovincialis*, *C. gigas* 600-mL glass beaker: *H. rufescens* (30-mL vial can also be used)
Test solution volume	10 mL: *S purpuratus*, *M. galloprovincialis*, *C. gigas*, *D. excentricus* 200 mL: *H. rufescens* (10 mL also used)
Number of larvae/mL test solution	5 to 10: *H. rufescens* 15 to 30: *M. galloprovincialis*, *C. gigas* 25: *S. purpuratus*
Number of replicates/treatment	Depends on objectives of test; usually 5 for *H. rufescens*, 4 for others
Dilution water	Uncontaminated 1 μm-filtered natural seawater/hypersaline brine
Test concentrations	Effluents: minimum 5 and a control Receiving waters: 100% and a control
Dilution factor	Effluents: ≥ 0.5 Receiving waters: 100%
Feeding	None
Aeration	None
Water quality	Temperature, pH, NH_3, salinity, and dissolved oxygen at start and end of test.
Test duration	48 hours: *M. galloprovincialis*, *H. rufescens*, *D. excentricus*, *C. gigas* 72 hours: *S. purpuratus*

Table 1 (continued). Test conditions for conducting a water column toxicity test for larval invertebrate development: *Haliotis rufescens, Strongylocentrotus purpuratus, Dendraster excentricus, Mytilus galloprovincialis,* or *Crassostrea gigas.* (Compiled from SWRCB 1996)

Endpoints	Normal larval development
Test Acceptability Criteria—Positive Control Performance	Abalone: significant effect at 56 µg/L Zn, %MSD < 20% Bivalve: % MSD < 25%
Test Acceptability Criteria—Negative Control Performance	Development: urchin/sand dollar ≥ 80%, abalone ≥ 80%, mussel ≥ 50%, oyster ≥ 70%; Percent normal: bivalve, 90% of survivors %MSD: urchin ≤ 20%

Table 2. Test conditions for conducting a water column toxicity test for fertilization success with *Strongylocentrotus purpuratus* or *Dendraster excentricus*.
(Compiled from SWRCB 1996)

Parameter	Conditions
Test type	Water column, static, non-renewal
Temperature	12ºC
Salinity	34‰
Light quality	Ambient laboratory light
Illuminance	10 to 20 μE/m^2/s (ambient laboratory levels)
Photoperiod	NA
Test chamber	16 x 100 or 16 x 125 mm glass test tube
Test solution volume	5 mL
Number of spawners	Pooled sperm and eggs from up to four individuals each
Number of egg and sperm cells per chamber	1,120 eggs; not more than 3,360,000 sperm
Number of replicates/treatment	4
Dilution water	Uncontaminated 1 μm-filtered natural seawater/hypersaline brine
Test concentrations	Effluents: minimum 5 and a control Receiving waters: 100% and a control
Dilution factor	Effluents: ≥ 0.5 Receiving waters: none or ≥ 0.5
Feeding	None
Aeration	None
Water quality	Temperature, pH, NH_3, salinity, and dissolved oxygen at start of test.
Endpoints	Fertilization of eggs
Test duration	40 minutes
Test Acceptability Criteria	1) Fertilization at NOEC > 80% 2) %MSD < 25% 3) Final sperm stock < 33,600,000/mL 4) Sperm stock > 5,600,000 not to exceed 2x target density, or high-density control fertilization 5% higher than low-density control

Embryo-larval Development Test Strengths, Limitations and Potential Confounding Factors

The primary strength of the WET toxicity test protocols currently listed in the U.S. EPA (1995a) West Coast Marine Toxicity Test Manual and the California Ocean Plan (SWRCB 2000) is that these are standardized methods that have been subjected to rigorous state and national selection criteria prior to their implementation as water quality regulatory tools. All procedures use ecologically relevant species indigenous to California waters, and these protocols have been demonstrated to be sensitive to a wide variety of toxicants. Repeated inter- and intralaboratory tests with reference toxicants and complex effluent samples have demonstrated that the precision of these procedures is comparable to analytical chemistry techniques (U.S. EPA 1995a). Although all of these protocols are considered to be appropriate for water quality assessments, the following discussion of the strengths and limitations of each protocol, and some of the confounding factors that may affect them, is intended to further guide their application in Ecological Risk Assessments.

Embryo-larval development tests using echinoids, gastropods and bivalve mollusks have been used in water and sediment quality assessments for decades. These tests are particularly useful for toxicity monitoring purposes because they require relatively short-term exposures ($\leq$ 96 h), yet incorporate sensitive, sublethal endpoints that represent critical life stages of ecologically important marine and estuarine species. Using the red abalone embryo-larval development test as an example, Hunt and Anderson (1989) demonstrated the ecological relevance of the larval development endpoint. In these experiments, zinc-exposed embryos that did not develop normally shaped veliger larval shells were shown to be incompetent to proceed to the next developmental stage, settlement and metamorphosis. Therefore, abalone embryos that do not develop normally will not enter the population. Woelke (1967) used *in situ* exposures with oyster (*C. gigas*) embryos to demonstrate receiving water toxicity in the vicinity of pulp and paper mills in Puget Sound, Washington, and showed a rapid elimination of ambient water toxicity when effluent discharges were stopped. No studies have been reported linking effluent or ambient water toxicity to marine ecosystem impacts. Examples of studies describing correlations between embryo development test results in sediment porewater

exposures and impacts on benthic community structure are discussed in the sections on porewater toxicity tests.

An additional strength of the embryo-larval development tests is that they are tolerant of Toxicity Identification Evaluation (TIE) procedures (U.S. EPA 1996). One potential consideration associated with the use of embryo-larval tests in water column toxicity assessments is their sensitivity to unionized ammonia toxicity.

Echinoid Fertilization Test Strengths, Limitations and Potential Confounding Factors

The sea urchin fertilization test was first developed for eastern Pacific species by Dinnel *et al.* (1983). This procedure has proven to be a sensitive indicator of effluent and ambient water toxicity. The fertilization test is among the most sensitive to certain chemicals, particularly metals, and is particularly useful as a screening test for large numbers of samples because it can be conducted quickly (*e.g.*, Bay *et al.* 1999). This attribute also makes the fertilization test useful for investigating toxicity of highly volatile or transient chemicals (*e.g.*, chlorine; Bay *et al.* 1993). A number of different echinoid species have been used for this test. On the Atlantic and Gulf Coasts the red urchin *Arabacia punctulata* is used. Although the purple sea urchin *Strongylocentrotus purpuratus* is the most commonly used species, a number of alternative echinoid species have been used on the West Coast (*S. franciscanus, S. droebachiensis, D. excentricus, L. pictus*). This protocol is amenable to TIE procedures and several studies have used the sea urchin fertilization test to identify causes of ambient toxicity (*e.g.*, sediment toxicity due to PCBs in New Bedford Harbor, Ho *et al.* 1997; storm water toxicity due to cationic metals in Southern California coastal waters, Bay *et al.* 1999). One other positive attribute of this protocol is the high tolerance of echinoid sperm to elevated unionized ammonia concentrations. This characteristic makes this test particularly useful in situations where ammonia may mask toxicity of other contaminants. Examples of studies describing correlations between fertilization test results using sediment porewaters and impacts on benthic community structure are discussed in the sections on porewater toxicity tests.

Bay *et al.* (1993) listed a number of limitations with this method. These authors noted toxicity artifacts associated with commercial sea salts and hypersaline brines with

tests conducted with *S. purpuratus* and *D. excentricus*, and also described the sensitivity of this test to pH extremes. These authors also discussed the occurrence of an unusually high rate of "false positive" toxicity results with this method when it has been used to assess ambient toxicity. False positive results occur when apparently non-toxic samples are identified as toxic. Bay *et al.* (in review) suggested that additional investigations on possible causes of unexplained toxicity in the echinoid fertilization test (*e.g.*, TIEs) should be conducted to clarify the significance of these events.

Marine Mysid Water Toxicity Tests

Mysid (*Holmesimysis costata* and Alternate Species) 96-h and 7-d Toxicity Tests

Compiled from SWRCB 1996, ASTM 2000a

Holmesimysis costata is a mysid crustacean found in the surface canopy of giant kelp beds off the Pacific coast, where it serves as an important food source for fish. These relatively short-lived, small crustaceans brood their developing young. Brooding adult females are collected from the kelp canopy and reared in seawater tanks, where fully developed live juveniles are released. These juveniles are collected and raised to an age of three or four days, at which time they are used in toxicity testing.

The *H. costata* tests are conducted in covered 600-mL or 1-liter jars containing 200 mL of test solution. Test solutions can consist of marine samples, salted fresh or estuarine samples, seawater/saltwater controls, and reference toxicant controls. Five test organisms are added to each test container, and are fed twice daily with newly hatched *Artemia* (brine shrimp). Test solutions are replicated five times each, and are renewed after 48 and 96 hours. Daily observations of survival are made so that 96-h and 7-d mortality endpoints can be determined. At the end of seven days, the test organisms are dried and weighed on a microbalance for determination of a growth endpoint.

Other species are often used in mysid short-term tests. *Neomysis mercedis*, another west-coast mysid, can be substituted in tests run at certain temperatures. Its optimum test salinity range is 1 to 3‰, but it can survive in the wild in salinities up to 18‰. *Americamysis* (*Mysidopsis*) *bahia*, a Gulf-coast mysid, is often used in chronic testing because of its short life cycle, and is generally tested in much warmer water at a salinity of 15 to 30‰.

Table 3. Test conditions for conducting a water column toxicity test with juvenile mysids: *Holmesimysis costata, Neomysis mercedis,* or *Americamysis* (*Mysidopsis*) *bahia*.
(Compiled from SWRCB 1996, ASTM 2000a)

Parameter	Conditions
Test type	Water column, static, renewal
Temperature	13 or 15ºC: *H. costata* 15 to 19ºC: *N. mercedis* 27°C: *A. bahia*
Salinity	34‰: *H. costata* 1 to 3‰: *N. mercedis* 15 to 30‰: *A. bahia*
Light quality	Ambient laboratory light
Illuminance	10 to 20 μE/m^2/s (ambient laboratory levels)
Photoperiod	16 light:8 dark
Test chamber	1-L glass beaker: mysids
Test solution volume	200 mL
Renewal and cleaning of test solutions	Renew 75% at 48 and 96 hours
Age of test organisms	3 to 4 days
Number of individuals/test chamber	5
Number of replicates/treatment	5
Dilution water	Uncontaminated 1 μm-filtered natural seawater/hypersaline brine
Test concentrations	Effluents: minimum 5 and a control Receiving waters: 100% and a control
Dilution factor	Effluents: ≥ 0.5 Receiving waters: none, or ≥ 0.5
Feeding	Newly hatched *Artemia* nauplii, twice daily
Aeration	None, unless dissolved oxygen is low
Water quality	Temperature, pH, NH_3, salinity, and dissolved oxygen at start and end of test, and at renewal
Test duration	4 or 7 days
Endpoints	Survival, growth
Test Acceptability Criteria—Positive Control Performance	Mysid survival %MSD < 40%; Mysid weight %MSD < 50% Mysid NOECs < 100 μg/L Zn
Test Acceptability Criteria—Negative Control Performance	Survival: mysid ≥ 75% Mysid weight ≥ 40 μg

Mysid Toxicity Test Strengths, Limitations and Potential Confounding Factors

Toxicity test protocols with mysid crustaceans have been used extensively in effluent and receiving water monitoring. Mysids represent a particularly important group for ambient monitoring because crustacea are sensitive to a variety of contaminants, including metals and metalloid compounds, and organochlorine, organophosphorous, and pyrethroid pesticides. The two species most commonly used in California are the kelp forest mysid *Holmesimysis costata* (U.S. EPA 1995a), and the estuarine species *Neomysis mercedis* (ASTM 2000a). *Neomysis mercedis* has been demonstrated to be useful in fresh and brackish water ambient monitoring studies where receiving water conductivities are beyond the range tolerated by other crustacea such as cladocerans (*e.g.*, Hunt *et al.* 1999, Finlayson *et al.* 1991).

TIE methods have been developed for *A. bahia* (U.S. EPA 1996), and *N. mercedis* (Hunt *et al.* 1999), and tolerances of *H. costata* to various TIE manipulations are now being assessed as part of the State Water Resources Control Board's Marine Bioassay Project.

The 7-d growth and survival test with *H. costata* may be limited by test organism availability because commercial suppliers have reported limited availability during winter (J. Hunt, personal communication). In addition, some researchers have reported difficulty meeting control performance using this species in 7-d tests (M. Swartz, Ogden International, personal communication). Mysid tests may be confounded by ionic concentrations above or below specific effect thresholds, particularly in certain effluents (*e.g.*, produced water and agricultural drain water; Ho and Caudle 1997). Pillard *et al.* (2000) developed models to predict the toxicity of elevated major ion concentrations and effects related to their deficiencies (K^+, Ca^{2+}, Mg^{2+}, Br^- SO_4^{2-}, HCO^{3-}, $B_4O_7^{2-}$).

Marine Fish Water Toxicity Tests

Topsmelt (*Atherinops affinis*) and Alternate Species 96-h and 7-d Toxicity Tests

Compiled from SWRCB 1996, ASTM 2000b

The topsmelt, *Atherinops affinis*, is a west-coast fish species commonly found in bays and estuaries during their reproductive season (summer). The adults are collected and reared by commercial labs that provide larvae for toxicity testing.

Atherinops affinis tests are conducted in covered 600-mL or 1-liter jars containing 200 mL of test solution. Test solutions can consist of marine samples, salted fresh or estuarine samples, seawater/saltwater controls, and reference toxicant controls. *Atherinops* can be tested at a broad range of salinities because of its greater tolerance to euryhaline conditions. Five test organisms are added to each test container, and are fed twice daily with *Artemia* (brine shrimp). Test solutions are replicated five times each, and are renewed daily. Tests are fed newly hatched *Artemia* nauplii twice daily. Daily observations of survival are made so that 96-h and 7-d mortality endpoints can be determined. At the end of seven days, the test organisms are dried and weighed for determination of a growth endpoint.

Other species are often used in fish short-term tests. *Menidia beryllina*, another atherinid fish species, can be substituted for *A. affinis* as a test organism. In addition, an embryo-larval development test, and a 7-d larval growth and survival test have been developed with the sheepshead minnow *Cyprinodon variegatus,* a species indigenous to the Atlantic and Gulf coasts (U.S. EPA 1994a).

Table 4. Test conditions for conducting a water column toxicity test with larval fishes: *Atherinops affinis* or *Menidia beryllina*.
(Compiled from SWRCB 1996, ASTM 2000b)

Parameter	Conditions
Test type	Water column, static, renewal
Temperature	20º C: fishes
Salinity	34‰: *M. beryllina* 5 to 34‰: *A. affinis*
Light quality	Ambient laboratory light
Illuminance	10 to 20 μE/m^2/s (ambient laboratory levels)
Photoperiod	16 light:8 dark
Test chamber	600-mL (or 1-L) glass beaker
Test solution volume	200 mL
Renewal and cleaning of test solutions	Renew maximum possible daily
Age of test organisms	9 to 15 days
Number of individuals/test chamber	5
Number of replicates/treatment	5
Dilution water	Uncontaminated 1 μm-filtered natural seawater/hypersaline brine
Test concentrations	Effluents: minimum 5 and a control Receiving waters: 100% and a control
Dilution factor	Effluents: ≥ 0.5 Receiving waters: none, or ≥ 0.5
Feeding	Newly hatched *Artemia* nauplii, twice daily (40 per fish)
Aeration	None, unless dissolved oxygen is low
Water quality	Temperature, pH, NH_3, salinity, and dissolved oxygen at start and end of test, and at renewal
Test duration	4 or 7 days
Endpoints	Survival, growth
Test Acceptability Criteria—Positive Control Performance	Fish Cu LC50 within 2 SD of control chart mean, and ≤ 205 μg/L; Fish survival %MSD < 25%; Fish weight %MSD < 50%
Test Acceptability Criteria—Negative Control Performance	Survival ≥ 80%; 9-day old mean weight > 0.85 mg

Fish Marine Toxicity Test Strengths, Limitations and Potential Confounding Factors

The topsmelt, *Atherinops affinis*, is one of the most ecologically important fish species in California estuaries, often representing the greatest fish biomass in these systems. The 7-d larval growth and survival protocol with this species is analogous to the test protocol for *Menidia beryllina* and other atherinid species, and was designed to be used in place of the *M. beryllina* protocol in West Coast testing. In addition to the larval growth and survival protocol with *A. affinis*, Anderson *et al.* (1991) developed a 12-d embryo-larval development test with this species. This test was not considered practical for routine effluent testing but may be appropriate in situations where teratogens are of particular concern. One of the strengths of this protocol is that topsmelt are a euryhaline species tolerant of a wide range of salinities. TIE methods have been developed for topsmelt and other atherinid larvae (U.S. EPA 1996)

Although topsmelt demonstrate comparable or greater sensitivity relative to other atherinid species (Middaugh and Anderson 1993), use of topsmelt in water quality assessments may be limited by lack of sensitivity relative to other fish and invertebrate species. Topsmelt larvae may be useful as an indicator of unionized ammonia toxicity in estuarine situations (*e.g.*, Nicely *et al.* 2000), and ancillary data suggest topsmelt larvae are particularly sensitive to low dissolved oxygen conditions (D. Middaugh, U.S. EPA, pers. comm.). Although studies have not been conducted with topsmelt, investigations using other atherinid species, and *Cyprinodon variegatus*, suggest that larval fish are sensitive to ionic imbalances, and this may confound results of tests with these species (Pillard *et al.* 2000). Anderson *et al.* (1995) found that, although topsmelt embryos and larvae are tolerant of salinities ranging from 5 – 35 ‰, experimental evidence suggests that larvae at lower salinities (≤ 17 ‰) may be more sensitive to contaminants due to osmotic stress.

Marine Algae Water Toxicity Tests

Giant Kelp (*Macrocystis pyrifera)* 48-h Spore Germination and Growth Toxicity Test

Compiled from U.S. EPA 1995a

Macrocystis pyrifera is a large marine alga that forms extensive forests in near-shore areas on the Pacific coast. These forests are structurally complex and provide habitat and food for numerous species. This kelp has a two-phase life cycle that alternates between the large, spore-forming stage (sporophyte) and the microscopic, gamete-producing stage (gametophyte).

Spore-producing fronds are collected from the base of wild plants. These sporophylls are subjected to cool, dry conditions, followed by immersion in seawater, resulting in spore release. Spores are collected, diluted to a known concentration, and inoculated into 200 mL of test solution in 600-mL containers, for a final density of 7500 spores per mL. Test solutions can consist of marine samples, salted fresh or estuarine samples, seawater/saltwater controls, and reference toxicant controls. The static, non-renewal test proceeds for two days, during which time the spores settle and germinate, developing into gametophytes. Two endpoints are measured: spore germination success and length of gametophyte germ tubes.

Table 5. Test conditions for conducting a water column toxicity test for spore germination and germ tube elongation: *Macrocystis pyrifera*.
(Compiled from SWRCB 1996)

Parameter	Conditions
Test type	Water column, static, non-renewal
Temperature	15ºC
Salinity	34‰
Light quality	Ambient laboratory light
Illuminance	50 μE/m^2/s
Photoperiod	16 light:8 dark
Test chamber	600-mL glass beaker (others can be used)
Test solution volume	200 mL
Spore density	7500/mL test solution
Number of replicates/treatment	5
Dilution water	Uncontaminated 1 μm-filtered natural seawater
Test concentrations	Effluents: minimum 5 and a control Receiving waters: 100% and a control
Dilution factor	Effluents: ≥ 0.5 Receiving waters: none or ≥ 0.5
Water quality	Temperature, pH, NH_3, salinity, and dissolved oxygen at start and end of test.
Test duration	48 hours
Endpoints	Germination and germ tube elongation
Test Acceptability Criteria—Positive Control Performance	Germ-tube NOEC < 35 μg/L Cu Both MSDs < 20%
Test Acceptability Criteria—Negative Control Performance	Germination: ≥ 70% Germ-tube length: ≥ 10 μm

Macrocystis pyrifera Toxicity Test Strengths, Limitations and Potential Confounding Factors

The 48-h test protocol using spores of the giant kelp, *Macrocystis pyrifera*, was developed to provide an algal toxicity test protocol for marine effluent monitoring. This protocol is among the most commonly used test protocols in California NPDES monitoring. Thursby *et al.* (1993) reviewed use of marine and estuarine micro- and macroalgae in toxicity testing and noted that one of the main reasons for including algal test protocols in water quality assessments is that, as primary producers, algae represent the foundation of aquatic food webs. Because of the prevalence of herbicides, fungicides, and other chemicals specifically designed to affect algae in many effluent and ambient samples, algal tests are a necessary component of water quality assessment programs. As part of the State Water Board's Marine Bioassay Project, both short-term chronic (48-h) and longer-term reproductive tests (21-d) were developed. Because of the practicalities involved, the 48-h test was developed for routine effluent testing, while the reproductive test was used to calibrate the relative sensitivity of the 48-h test. A number of studies demonstrated the ecological significance of the 48-h test endpoints with *M. pyrifera* spores. These experiments showed that toxicants that inhibited spore germination and growth also inhibited kelp reproduction (sporophyte production). TIE methods have been developed for this test (U.S. EPA 1996).

One potential confounding factor associated with the 48-h test with *M. pyrifera* spores has been reported. In this test, kelp reproductive blades called sporophylls are collected from the field the day before the test is initiated. Sporophylls are shipped to the testing laboratory, where they are then immersed in seawater to induce spore release. The time between sporophyll collection and spore release is typically < 24-h. Gully *et al.* (1999) found that sporophyll storage affected response of the spore germination endpoint in reference toxicant tests with copper. While these authors found no affect on the germ-tube growth endpoint, they suggested possible affects on the germination endpoint may confound interpretation of effluent tests with this protocol by increasing the relative sensitivity of the germ-tube growth endpoint. Affects of sporophyll storage on kelp spore energetics are the subject of current State Water Resources Control Board research. One other limitation of this protocol is that it may be less appropriate for testing estuarine

samples. *Macrocystis* is a coastal species restricted to rocky subtidal areas. An alternative test for studies concerning algal toxicity in estuarine habitats has been reported by Hooten and Carr (1998). This test is analogous to the protocol for *M. pyrifera* but uses zoospores of the estuarine alga *Ulva fasciata*. These authors evaluated this test for sediment porewater testing and suggest that because *Ulva fasciata* spores are relatively sensitive to a number of toxicants and are tolerant of unionized ammonia, this test may be useful in situations where elevated NH_3 is a potential confounding factor.

Relative Contaminant Sensitivities of Marine Water Column Toxicity Tests

The relative sensitivities of the seven California Ocean Plan Whole Effluent Toxicity (WET) test protocols using *Haliotis rufescens, Holmesimysis costata, Atherinops affinis, Macrocystis pyrifera, Mytilus galloprovincialis,* and *Strongylocentrotus purpuratus* (fertilization and development tests) can be compared by assessing responses in water-only exposures (*e.g.,* reference toxicant tests), and by comparing side-by-side tests of effluents or ambient samples. Results of water-only exposures were compiled from those reported in the literature (*e.g.,* ECOTOX database and other sources) and in some cases, from unreported experiments conducted at testing laboratories contacted by UC Davis. Results of these experiments are generally reported as 96-h LC or EC50 values and can be found in the US EPA ECOTOX database (www.epa.gov/ecotox).

A summary of dose-response data for selected chemicals shows that sensitivity varies between methods, and there is no consistent pattern. Embryo-larval development tests with bivalves, sea urchins, and abalone, and fertilization tests with sea urchins often show greater sensitivity to copper and zinc. Topsmelt (*A. affinis*) larvae are much less sensitive to copper, while kelp spores (*M. pyrifera*) demonstrate moderate sensitivity to this metal relative to the invertebrate embryo-larval protocols. Based on available data, tests with mysids appear to be more sensitive to cadmium than the other marine protocols. Comparisons between these protocols are limited by a lack of dose-response data for pesticides and other organic compounds. Available data indicate that mysids are particularly sensitive to certain pesticides and biocides. Mysids (*H. costata* and/or *A.*

bahia) were quite sensitive to the organochlorine pesticide endosulfan, and the organophosphate pesticide diazinon. Topsmelt larvae were also relatively sensitive to endosulfan. Based on comparisons of data from the US EPA ECOTOX database, mysid neonates and abalone embryos demonstrated comparable sensitivity to pentachlorophenol. Of the tests for which data exist, mysids were far more sensitive to the biocide sodium azide, and were among the most sensitive group to PCBs (Arochlor 1254). Embryo-larval tests with bivalves and the fertilization test with sand dollars (echinoids) were also more sensitive to tributyltin (TBT) than tests with kelp spores. Embryo-larval tests with sea urchins, bivalves, and abalones, and larval tests with topsmelt demonstrate greater sensitivity to unionized ammonia than tests with mysids or fertilization tests with sea urchins.

A number of protocol comparisons were conducted using complex effluent samples as part of the State Water Board's Marine Bioassay Project. Hunt *et al.* (1989) assessed toxicity of two sewage treatment plant effluents (Plant A and B) using three tests: the 48-h kelp germination and germ-tube growth test (*M. pyrifera*), the 48-h abalone development test (*H. rufescens*), and the 96-h mysid survival test (*H. costata*). Separate samples of both effluents were tested at two different times. Sensitivities (greatest to least sensitive) to the first sample of Plant A were as follows: abalone development > kelp germ-tube growth > mysid survival > kelp germination. Sensitivities to the second sample of Plant A were as follows: kelp germ-tube growth > mysid survival > kelp germination = abalone development. Sensitivities to the first sample of Plant B were as follows: abalone development = mysid survival > kelp germ-tube growth > kelp germination. Sensitivities to the second sample of Plant B were as follows: abalone development > mysid survival > kelp germ-tube growth > kelp germination. These tests were not designed to investigate causes of effluent toxicity, but the results suggest that causes varied over time at each effluent source. Hunt *et al.* (1991) conducted similar comparisons of four sewage effluents using the kelp and abalone protocols. This study indicated that kelp germ-tube growth was a more sensitive indicator of toxicity in tests conducted on 3 of the 4 effluents. The kelp germination endpoint was not assessed in these experiments.

Emphasis on water quality assessments has recently shifted from point source to non-point source pollution impacts. Some of the marine water column protocols discussed above have been used to assess ambient water toxicity. Anderson *et al.* (1998a) assessed water column toxicity using samples collected from Moss Landing Harbor, Monterey Harbor, and 3 sites in San Francisco Bay, as part of a study to evaluate California Ocean Plan protocols for use in ambient water assessments. In most cases, samples were collected from sites that had previously been characterized as toxic hotspots because of sediment contamination and observed biological effects. Laboratory tests of field-collected water were conducted using sea urchin and bivalve embryos (*S. purpuratus* and *M. galloprovincialis*), and mysid neonates (*H. costata*). Although all sites had previously exhibited sediment toxicity, water column samples collected from these sites were not toxic during this survey.

In addition to laboratory tests of ambient water samples, these authors used bivalve embryo development to assess toxicity of sediment samples from Monterey Harbor using laboratory and *in situ* exposures. Bivalve embryos were exposed to sediment samples at the sediment-water interface using a polycarbonate exposure chamber. Sediments from one Monterey Harbor station were toxic to bivalve embryos in both laboratory and *in situ* exposures, although the magnitude of response was considerably greater in the laboratory-exposed animals. Toxicity at this site was presumably due to chemicals fluxed from the sediment into the overlying water. Differences between the laboratory and *in situ* exposures may have been due to lower pH in the laboratory exposure chambers. These authors concluded that these tests were amenable to *in situ* exposures and laboratory testing of ambient samples with minimal modification.

Bay *et al.* (1999) used the sea urchin (*S. purpuratus*) fertilization test to assess toxicity of storm water samples entering marine waters at two sites in Southern California. In this study, ocean surface water impacted by storm water from Ballona Creek was consistently toxic to sea urchin sperm, and toxicity decreased with distance from the creek input. TIEs conducted on selected samples showed that sea urchin fertilization rates increased with the addition of EDTA, indicating toxicity due to divalent cations. These authors concluded the sea urchin fertilization protocol was a sensitive

indicator of storm water toxicity, and was particularly useful as a short-term screening tool for tracking the spatial extent of storm water toxicity plumes entering marine waters.

Application of Marine Water Column Toxicity Tests

Compiled from U.S. EPA 1995a

The protocols listed above have been used extensively for effluent toxicity monitoring, and to a lesser extent in ambient water monitoring. In 1995, the Society of Environmental Toxicology and Chemistry (SETAC) convened a workshop in Pellston, Michigan, to evaluate current methods for using Whole Effluent Toxicity tests in effluent and ambient water quality assessments. This workshop consisted of experts from government, industry, and academia who were experienced in issues concerning the use of WET tests for these applications. The consensus of the workshop participants was that these test protocols are technically sound when conducted according to U.S. EPA methods. Although the workshop participants concluded that these tests provide useful information on the potential for effluents to impact receiving waters, the application of these tests for marine and estuarine ambient water toxicity monitoring has not been as thoroughly evaluated as in freshwater systems (Grothe *et al.* 1996). The workshop proceedings identified several areas where more research is needed. Schimmel and Thursby (1996) noted that for a variety of reasons, no studies have been conducted to link ambient toxicity in marine or estuarine receiving waters with impacts on water column or benthic communities in those systems. The relationship between ambient toxicity and receiving system ecological impacts are more difficult to ascertain in these systems because of the complex biotic and abiotic factors that may interact with chemical stressors in these systems. The workshop participants concluded that water column toxicity tests such as standardized WET tests should be used in concert with biological assessments and chemical analyses for integrated decision-making (as described at the end of this review).

For effluent testing purposes, the California Ocean Plan (SWRCB 2000) recommends a minimum of three of the test protocols listed above be used to screen effluent samples for toxicity. If possible, the test species shall include a fish, invertebrate, and an aquatic plant, because these groups may respond differently to

different classes of toxicants. The practice of including a suite of test species representing different phyla and groups also applies to ambient toxicity studies (U.S. EPA 1991a). Because this section of the review is concerned with ambient water column toxicity testing as part of Ecological Risk Assessments, the following discussion provides guidance for using these protocols for this application.

As discussed above, state and federal guidance on application of water column toxicity tests for ambient water quality monitoring suggests that a toxicity screening phase be conducted with a minimum of three species representing a variety of groups including invertebrates, fish and plants. Subsequent testing can then be done with the most sensitive species. Because protocol sensitivities vary both between and within these groups, selection of appropriate protocols for use in effect characterizations in ERAs depends on the chemicals of concern identified in the problem formulation stage of the risk assessment. For example, relative to mysids, embryo-larval development tests with echinoids and mollusks, and fertilization tests with echinoids are not particularly sensitive to cadmium; therefore, screening with an invertebrate test other than mysids might underestimate ecological risk posed by this metal. Conversely, echinoid and molluscan embryo-larval development tests and fertilization tests are the most sensitive of the standardized protocols to copper and zinc. In situations where these metals are the primary chemicals of concern, these protocols would be more appropriate. Note that although the embryo-larval development (*M. galloprovincialis, S. purpuratus, H. rufescens*) tests are often grouped together because they incorporate similar endpoints, these protocols may not respond similarly to all toxicants. For example, Phillips *et al.* (in review) found considerable difference between mussel and sea urchin embryos in response to cadmium, copper, zinc, and nickel. Similar differences between sea urchin and bivalve embryos have been reported by others (*e.g.*, Gries 1998).

Like other crustacea (*e.g.*, amphipods), mysids are also sensitive to many general biocides (*e.g.*, sodium azide, pentachlorophenol) and pesticides, particularly organochlorine and organophosphorous pesticides. Mysids are also relatively sensitive to other organochlorine compounds, such as PCBs. The test using *Holmesimysis costata* or an alternative species (*N. mercedis or A. bahia*) would be appropriate when these are the primary contaminants of concern. The 7-d growth and survival test using *H. costata* does

not include a reproductive endpoint. If reproductive effects on mysids or other crustacea are of concern in a particular risk assessment, testing with the Gulf Coast species (*A. bahia*) is an appropriate surrogate. A similar west coast species, *Mysidopsis intii* (Langdon *et al.* 1996), has also been used in tests designed to incorporate reproductive endpoints (*i.e.*, fecundity). Given their sensitivity, mysids, sea urchin fertilization tests, and embryo-larval development tests with some species are also appropriate for risk assessments associated with some organochlorine pesticides (*e.g.*, DDT) and metalloid compounds (*e.g.*, TBT). As discussed previously, the sea urchin fertilization test (*S. purpuratus* or *D. excentricus*) is sensitive to a wide variety of toxicants, and is particularly useful for screening highly volatile or transient chemicals (*e.g.*, chlorine, storm water). Spores of the giant kelp (*M. pyrifera*) have also been shown to be quite sensitive to chlorine (T. Dean, Coastal Resource Associates, personal communication). Tests using spores of marine and estuarine algae (*M. pyrifera, U. fasciata*) are also applicable for risks associated with fungicides and herbicides.

In some cases there is insufficient data to determine relative sensitivity of marine water column toxicity test protocols to certain contaminant classes. For example, few comparative studies have been conducted to assess the relative sensitivity of these protocols to PAHs. Ancillary research indicates that, because of their apparent sensitivity, protocols using larval marine fish are appropriate for risk assessments where petroleum hydrocarbons are of concern. For example, Schiff *et al.* (1992) found that silverside larvae (*M. beryllina*) were among the most sensitive of 5 protocols tested with produced water (*S. purpuratus* fertilization>*M. beryllina* larval survival>*A. bahia* neonate survival> Microtox>*N. arenaceodentata* survival). In tests assessing the interactive effects of chemical dispersants and oil, Singer *et al.* (1998) found that topsmelt larvae (*A. affinis*) were sometimes the most sensitive species to the water-accommodated fraction (WAF) of Prudhoe Bay crude oil, compared to abalone embryos (*H. rufescens*), and mysid neonates (*H. costata*). When dispersants were used to chemically enhance the preparation of the Prudhoe Bay WAF, mysids were more sensitive that both other species. Fish larvae are also appropriate in situations where unionized ammonia is a chemical of concern.

Many of these protocols are sensitive to non-contaminant factors and naturally occurring compounds that may confound interpretation of toxicity test results. For example, mysids, fish larvae, and in some cases sea urchin sperm may be affected by ion concentrations above or below effect thresholds (Bay *et al.* in review, Pillard *et al.* 2000). In ERAs where ambient waters may be influenced by produced water, agricultural drain water, or other sources that may increase ion concentrations, these constituents should be measured and compared to established effect models (K^+, Ca^{2+}, Mg^{2+}, Br^- SO_4^{2-}, HCO^{3-}, $B_4O_7^{2-}$). In addition, many of these test protocols are sensitive to elevated unionized ammonia. Because all of these protocols are amenable to Toxicity Identification Evaluation procedures, these procedures are useful to confirm causes of toxicity, particularly when non-contaminant factors affect the results.

As discussed previously, because of the variable sensitivities of these protocols to contaminants, US EPA recommends testing with multiple protocols representing a variety of phyla and groups. This is especially important where ambient waters may be impacted by complex chemical mixtures. Schimmel *et al.* (1989) assessed the toxicity of 7 different effluents and their receiving waters using 5 different Atlantic coast toxicity test protocols (*C. parvula, A. bahia, A. punctulata, M. beryllina, C. variegatus*). Sensitivity to effluents and receiving waters varied between protocols and no one protocol was the most sensitive to every effluent or receiving water sample. In addition to using multiple species in standardized protocols, additional endpoints may be assessed with many of these protocols to provide ancillary information regarding ecological risk. For example, cytogenetic endpoints have been assessed with sea urchin sperm and embryos, and with fish embryos and larvae (Anderson *et al.* 1994, Kocan 1996). Use of water column toxicity tests in integrated studies incorporating biological assessments, sediment toxicity tests and where appropriate, bioconcentration and bioaccumulation studies, is discussed below.

MARINE WATER COLUMN BIOCONCENTRATION TESTS

Bioconcentration Tests with Bivalves (Mytilus galloprovincialis, Crassostrea gigas, Cyprinodon variegatus) and Fish (Fundulus parvipinnis, Cymatogaster aggregata)
Compiled from ASTM 2000c

Bioconcentration tests are laboratory experiments designed to obtain information concerning the ability of aquatic species to accumulate chemicals directly from water. They are distinguished from bioaccumulation experiments, which are designed to consider all uptake pathways, including food. Bioconcentration data provide information to allow prediction of concentrations of test material likely to occur in aquatic organisms in field situations, and allow comparisons between species regarding their potential to accumulate chemicals. These procedures are designed for calculation of bioconcentration factors (BCFs) for specific chemicals of concern. A BCF is defined as the quotient, during the uptake phase of a test, of the tissue concentration of a test material divided by the average exposure concentration in water. Exposure durations vary depending on the time it takes a specific chemical to reach tissue steady-state; tests are not conducted for longer than 28 days. A number of species have been recommended for bioconcentration tests by ASTM (2000c). The species listed above may be more appropriate than other species for use in ERA applications in California waters.

Bioconcentration tests are conducted in aquaria under flow-through conditions using an appropriate metering system. No food is added in tests with bivalves; these tests use higher flow rates of unfiltered seawater to provide some natural planktonic food. When fish are used, test organisms are free. This food is chemically characterized prior to use in tests to ensure that it is contaminant free. Tests are conducted with sexually immature individuals such as juvenile fish or smaller bivalves (40 – 60 mm long), to limit the confounding effects of gonadal tissue differentiation.

In these experiments, animals exposed to a given concentration of a test material are compared to a control group exposed to dilution water. Treatment animals are exposed during an uptake phase until a tissue steady-state is achieved, or until 28-d is reached. The criterion for steady-state requires that there be no difference between three sets of BCFs taken at appropriate intervals. Animals are exposed during the depuration

phase to dilution water. During both phases of the test, representative organisms and water samples are periodically removed and analyzed for the test material. Apparent steady-state and projected steady-state BCFs and uptake and depuration rate constants are calculated from the measured concentrations of test material in tissue and water samples.

Sampling schedules for both uptake and depuration phases depend on the time it takes a test material to reach steady-state, and this time is estimated from octanol-water partition coefficients for particular compounds (Log K_{ow}). Minimum organism sampling schedules for uptake and depuration sampling phases based on representative chemical Log K_{ow} are provided in ASTM 2000c. For bioconcentration tests concerned with organic chemicals, concentrations of lipids are also measured in control and treatment test organisms at the beginning and end of the uptake and depuration phases of the test.

Marine Water Column Bioconcentration Test Strengths, Limitations and Potential Confounding Factors

See Marine Bioconcentration/Bioaccumulation Test Strengths, Limitations and Potential Confounding Factors.

MARINE WHOLE SEDIMENT TOXICITY TESTS

Marine Amphipod Sediment Toxicity Tests

Amphipod 10-d Survival Toxicity Test

Compiled from ASTM 2000e, U.S. EPA 1994b

Amphipod crustaceans are ecologically important members of benthic infaunal communities and are a primary food resource for a number of marine invertebrate, fish, and bird species world-wide. In general, crustacea are among the most sensitive members of benthic communities to anthropogenic disturbance, including pollution. The 10-day solid-phase amphipod survival toxicity test protocol is appropriate for a number of amphipod species. The following is a general description of the acute (10-d) toxicity test protocol as it pertains to all accepted species. Specific information on each species is provided in the succeeding discussion.

Except for *Leptocheirus plumulosus*, adult amphipods used for marine sediment toxicity assessments are usually provided by commercial suppliers from field-collected populations. Once shipped to testing laboratories, amphipods are acclimated to test salinity (≤5‰ change per day) and temperature (≤3ºC change per day). Test animals are then held for an additional 48 hours prior to inoculation into test containers. Test containers are one-liter glass beakers or jars containing 2 cm of homogenized sediment and filled with control seawater adjusted to the appropriate salinity. Two species, *Ampelisca* and *Leptocheirus* are routinely fed a combination of ground flake food and algae during the holding period (*Rhepoxynius, Eohaustorius* and *Grandidierella* are not fed). Test sediment and overlying water are allowed to equilibrate for 24 hours, after which 20 amphipods are placed in each beaker. Test chambers are aerated gently, and overlying water is generally not renewed, except in situations where test sediments have unionized ammonia concentrations above protocol thresholds. In these cases it is sometimes necessary to purge ammonia by renewing overlying water before amphipods are added to the beakers. Five laboratory replicates of each sample are tested for ten days. A negative-control sediment consisting of five laboratory replicates of home sediment (sediment from the amphipod collection site) is included with each sediment test. After ten days, the sediments are sieved to recover the test animals, and the number of survivors is recorded for each replicate. An alternative sublethal endpoint is described in the U.S. EPA (1994b) manual. The ability of surviving amphipods to rebury in clean control sediment can be used to calculate *effective mortality*. For this endpoint, surviving amphipods are collected at the end of the 10-d exposure and placed in separate replicate beakers containing control sediment, and having the same salinity as the test sediments. Those animals that fail to rebury after one hour are combined with the count of dead animals in a separate *effective mortality* category. This is an optional endpoint and has been used for *R. abronius, E. estuarius,* and *L. plumulosus,* but is not often reported in the literature. Recommended test conditions for all species are listed in Table 6. For the test to be acceptable, survival at 10-d must equal or exceed 90% for all five amphipod species in the negative control (home) sediment. Additional requirements are listed in Table 7.

Marine Amphipod Toxicity Test Strengths, Limitations and Potential Confounding Factors

The amphipod 10-d survival protocol is currently the most commonly used toxicity test for assessing marine sediment quality. Standardized peer-reviewed procedures are available for all species, and in many regards, this protocol is considered to be the benchmark test for marine sediment toxicology. Although this is an acute toxicity test procedure, amphipods are among the most sensitive infaunal groups to contaminants (Mearns and Word 1982), and are appropriate indicators of pollution. Because amphipods are exposed to contaminants via ingestion of sediment particles and dermal uptake through porewater exposure, they are useful for assessing effects of both hydrophobic and more water-soluble contaminants. Although all species for which the 10-d protocol has been developed are considered to be appropriate for sediment quality assessments, the discussion of the strengths and limitations of each species, and some of the confounding factors that may affect them, is intended to further guide their application in Ecological Risk Assessments.

Rhepoxynius abronius Toxicity Test

Rhepoxynius abronius is a phoxocephalid species that occurs in clean, fine-sandy sediments along the west coast of Northern America from central California to Puget Sound, Washington. This is a free burrowing carnivorous species that also ingests organic material, and is the amphipod species for which the 10-d survival protocol was originally developed by Swartz *et al.* (1979). A marine species, *R. abronius* is the least tolerant of all the currently available amphipods to low salinity sediments ($\geq$ 25 ‰). *Rhepoxynius* is also less tolerant of very fine-grained sediments and is not recommended for testing sediments having greater than 90% silt/clay (U.S. EPA 1994b). In situations where sample grain size distributions are unknown prior to testing, EPA (U.S. EPA 1994b) recommends including silt/clay reference sediments having particle size distributions that bracket test sediments. Numerous studies have shown that this species is sensitive to trace metal and organic contaminants. In addition, a number of studies have demonstrated that mortality in laboratory tests with *Rhepoxynius* is correlated with

declines in densities of amphipods and other infaunal groups in samples collected from the same stations.

Rhepoxynius abronius Toxicity Test Strengths, Limitations and Potential Confounding Factors

The 10-d survival test with the amphipod *Rhepoxynius abronius* has been used extensively in sediment toxicology studies because it was the species for which this protocol was originally developed. A considerable amount of information is therefore available for this species. Much of the published literature describes the usefulness of *R. abronius* in identifying contaminated sediments and the ecological relevance of laboratory toxicity tests with this species as an indicator of benthic community effects (Swartz *et al.* 1982, Swartz *et al.* 1985, Swartz *et al.* 1986, Swartz *et al.* 1989). It has been subjected to interlaboratory testing, and interlaboratory precision among five participating laboratories was within acceptable limits (Mearns *et al.* 1996).

One limitation with *R. abronius* is that it has not been demonstrated to be amenable to long-term laboratory culture; therefore, no chronic toxicity test protocol has been developed with this species. In situations where sediments are contaminated by moderate concentrations of contaminants, or are dominated by chemicals not considered to be acutely toxic, the 10-d test with this species may not detect potential for impacts. In addition, *R. abronius* naturally occurs in sandy marine habitats and does not tolerate silty/clay sediments, presumably because they impair gill function. DeWitt *et al.* (1988) found that survival of *R. abronius* may decline in fine-grained sediments, and developed regression procedures to account for grain size effects with this species. EPA (U.S. EPA 1994b) recommends not using *R. abronius* for assessing sediments having greater than 90% fines. Because it is a marine species, *R. abronius* is also not appropriate for testing estuarine sediments (*i.e.*, those with interstitial water salinities < 25 ‰). Knezovich *et al.* (1996) found *R. abronius* to be relatively sensitive to hydrogen sulfide, and Kohn *et al.* (1996) found this species to be sensitive to unionized ammonia. It is recommended that both of these sediment constituents be routinely measured (Phillips *et al.* 1997) as part of all sediment quality assessments. Other potential factors that may affect this and other sediment test species are listed in Table 8.

Eohaustorius estuarius Toxicity Test

Eohaustorius estuarius is a haustoriid species that occurs in sandy sediments in mid-intertidal to shallow subtidal habitats from central California to British Columbia. This amphipod is a free-burrowing detritivore and is presumably exposed to sediment contaminants via dermal uptake (respiration in porewater) and particle ingestion. *E. estuarius* is a euryhaline estuarine species, and is highly tolerant of a wide range of temperatures (~ 5 - 21ºC) and salinities (0 – 34 ‰). Although *E. estuarius* is considered to be less susceptible to fine-grained sediments, several studies have indicated negative correlations between survival and percent clay distributions with this species (U.S. EPA 1994b). As with *Rhepoxynius*, EPA (U.S. EPA 1994b) suggests including silt/clay reference sediments having particle size distributions that bracket test sediments in situations where sample grain-size distributions are unknown prior to testing with *E. estuarius*. The toxicity test with *E. estuarius* was first reported by DeWitt *et al.* (1989) who, in addition to studying salinity and grain size tolerance, found this species to have sensitivity comparable to that of *Rhepoxynius abronius* to fluoranthene-spiked and field-collected sediments. *Eohaustorius* has been used extensively in sediment monitoring and assessment programs in California (Fairey *et al.* 1998, Anderson *et al.* 2001a, Hunt *et al.* 2001 a and b).

Eohaustorius estuarius Toxicity Test Strengths, Limitations and Potential Confounding Factors

Eohaustorius estuarius has not been used as extensively as *R. abronius* or *A. abdita*, although its use is increasing in west coast sediment assessment studies. It has been subjected to interlaboratory testing and the protocol using this species was determined to demonstrate acceptable precision (Schlekat *et al.* 1995), and has been demonstrated to have comparable sensitivity to the other commonly tested amphipod species in comparisons using field sediments. *E. estuarius* has become the primary testing species in several recent regional and statewide monitoring programs in California (Thompson *et al.* 1999, Hunt *et al.* 2001a, Anderson *et al.* 2001a, Bay *et al.* 2000). In addition, *E. estuarius* is commonly used as a dredge material monitoring species in

Puget Sound (*e.g.*, Gries 2000), and is used in all testing of estuarine and fine-grained sediments in Canadian sediment quality assessment programs **(**K. Doe, Environment Canada, personal communication).

Anderson *et al.* (2001a) found that survival of *E. estuarius and R. abronius* in laboratory tests correlated with a number of benthic community metrics in field-collected sediments from Los Angeles Harbor, including the total number of fauna, and the number of crustacean species present in Los Angeles Harbor sediments. Because this is an estuarine species, it is tolerant of a wide range of salinities (0 – 34 ‰). It is also more tolerant of elevated NH_3 (Kohn *et al.* 1994) and H_2S (Knezovich *et al.* 1996) than some of the other commonly used species (Table 8).

This species has not been demonstrated to be amenable to long-term laboratory culture, although recent research into its adaptability to continuous culture is currently being conducted (S. Kellman, Aquatic Biosystems, personal communication). No chronic toxicity test protocol has been developed with this species. In situations where sediments are contaminated by moderate concentrations of contaminants, or are dominated by chemicals not considered to be acutely toxic, the 10-d test with this species may not detect potential for impacts.

There is some evidence to suggest survival of *E. estuarius* is negatively correlated with percent clay in field sediments (Hunt *et al.* in press), although DeWitt *et al.* (1989) found this species is tolerant of fine-grained sediments. EPA (U.S. EPA 1994b) recommends including reference sediments having grain-size distributions similar to the field sediments under consideration if *E. estuarius* is used. Other potential factors that may effects this and other species in sediment and porewater exposures are listed in Table 8.

Ampelisca abdita Toxicity Test

Ampelisca abdita is a euryhaline ampeliscid amphipod species that ranges from central Maine to south-central Florida, and has also been introduced to San Francisco Bay, California. This species is tolerant of a wide range of grain sizes (>10% silt/clay) and salinities (>10‰ to 34‰). Because it requires fine sediments for tube building, *Ampelisca* is less tolerant of very coarse-grained sediments. *Ampelisca* is a surface

feeding detritivore and constructs a tube that is partially permeable to porewater. It is exposed to contaminants via surface particle ingestion, and dermal uptake (*i.e.*, overlying water and porewater respiration). Robson (1990) used dye studies to demonstrate that while feeding *Ampelisca* are exposed primarily to overlying water, and when positioned at the base of the tube, they are exposed to porewater. This author concluded that porewater exposure may be proportional to the amount of time this species is not feeding. This species is routinely tested at 20ºC in 28‰ overlying water. *Ampelisca abdita* has been shown to be relatively sensitive to a variety of anthropogenic materials in a number of studies. *Ampelisca abdita* is the species most commonly used as the solid-phase sediment testing species in national sediment monitoring and assessment programs (*e.g.*, Environmental Monitoring and Assessment Program (EMAP), National Status and Trends Program (NS&T).

Ampelisca abdita Toxicity Test Strengths, Limitations and Potential Confounding Factors

Ampelisca abdita has also been tested extensively in Atlantic and Gulf coast sediment studies. *A. abdita* is the primary species used in two national sediment monitoring programs (EMAP, U.S. EPA 1995b; National Status and Trends program, NOAA, Long *et al.* 1998). It has been subjected to interlaboratory testing and the protocol using this species was determined to demonstrate acceptable precision (Schlekat *et al.* 1995). *A. abdita* is an estuarine species and is tolerant of a wide range of salinities. A considerable amount of information is available correlating low survival of this and other amphipod species to elevated bulk-phase sediment concentrations (Long *et al.* 1995, Long *et al.* 1998, Long *et al.* 2000). Long *et al.* (2001) examined the relationship between amphipod survival in 10-d toxicity tests using a variety of amphipod species and benthic community metrics. In a review of the combined EMAP data sets from the Atlantic and Gulf of Mexico coasts, these authors found that 72% of the samples had benthic conditions classified as degraded when survival of *A. abdita* was less than 80%, and 84% of the samples had degraded benthos when *A. abdita* survival was less than 60%.

Although some research has been devoted to developing a chronic protocol for this species (Scott and Redmond 1989, Redmond *et al.* 1994), a standardized chronic test protocol has not been completed. This is partly due to difficulties in achieving

continuous reproduction of *A. abdita* under laboratory conditions (M. Redmond personal communication).

A. abdita requires relatively fine-grained sediments for tube construction, so toxicity investigations of coarser sediments with this species may be confounded by grain size effects. EPA (U.S. EPA 1994b) recommends including reference sediments having grain-size distributions similar to the field sediments under consideration if *A. abdita* is used.

This species is sensitive to unionized ammonia, and EPA has established application limits for this constituent for sediment testing purposes (Unionized ammonia No Observed Effect Concentration = 0.4 mg/L; U.S. EPA 1994b). No information on toxicity of H_2S is available for *A. abdita*. Other potential factors that may effects this and other species in sediment and porewater exposures are listed in Table 8.

Leptocheirus plumulosus Toxicity Test

Leptocheirus plumulosus is an aorid amphipod species that is distributed subtidally in estuarine rivers and embayments along the east coast of the United States from Cape Cod, Massachusetts to northern Florida. *Leptocheirus* builds U-shaped burrows in sediments ranging from fine sand to silty clay and tolerates salinities from near 0 to 33 ‰. McGee *et al.* (1999) conducted sediment quality assessments in Baltimore Harbor that included chemical analyses, characterizations of benthic community structure, and 10-d toxicity tests using *Leptocheirus*. These authors found negative correlations between amphipod survival in laboratory exposures and bulk sediment contamination, and a strong positive correlation between *Leptocheirus* survival in laboratory exposures and the density of *Leptocheirus* in field samples.

Leptocheirus are more amenable to laboratory culture than the other commonly used marine amphipod species; therefore, known-age animals may be obtained for toxicity testing. This attribute, combined with a relatively short generation time (life span $\geq$ 7 weeks), has allowed for development of a 28-d toxicity test with *Leptocheirus* that incorporates sublethal endpoints (survival, growth, and reproduction). Initial comparisons between the 10-d and 28 d tests have suggested that the chronic test is not always more sensitive than the 10-d test (Farrar *et al.* 1999). These authors suggested

that the relative sensitivity of the 10-d and 28 d tests is compound specific and is also influenced by bioassay design (*e.g.*, feeding, water exchange).

Leptocheirus plumulosus Toxicity Test Strengths, Limitations and Potential Confounding Factors

Leptocheirus plumulosus is being used increasingly in Atlantic coast sediment quality assessment programs because it has many desirable attributes. This is an estuarine species that has been demonstrated to be highly tolerant of a wide range of salinities and sediment grain sizes (ASTM 2000e and references therein). McGee *et al.* (1999) found negative correlations between amphipod survival in laboratory exposures and bulk sediment contamination in Baltimore Harbor sediment, and a strong positive correlation between *Leptocheirus* survival in laboratory exposures and the density of *Leptocheirus* in field samples. It has been subjected to interlaboratory testing and the protocol using this species was determined to demonstrate acceptable precision (Schlekat *et al.* 1995). Because this species is amenable to laboratory culture and has a relatively short generation time, it has allowed development of a chronic toxicity test protocol that incorporates a number of sublethal endpoints including survival, growth, and fecundity.

Leptocheirus plumulosus is an east coast species and its appropriateness for predicting benthic impacts in other coastal systems has not yet been investigated. It is also relatively sensitive to unionized ammonia; no information on toxicity of H_2S is available for this species. Other potential factors that may affect this species are listed in Table 8.

Grandidierella japonica Toxicity Test

Grandidierella japonica is a corophiid amphipod species that was accidentally introduced from Japan into San Francisco Bay in the late 1960's. *Grandidierella japonica* was later found in some southern California bays. This species builds a U-shaped tube in sediments ranging from coarse sand to silty clay. A standardized 10-d protocol has been developed for this species (ASTM 1996). *Grandidierella* is not used as commonly as the other amphipod species for which standardized marine testing procedures are available partly because of concerns this species may adapt physiologically to contaminated sediments (*e.g.*, Swartz *et al.* 1994, Lamberson *et al.*

1996, Anderson *et al.* 2001a). Tests with reference toxicants and field collected sediments indicate the sensitivity of this species is comparable to other amphipods, provided test organisms are not collected from pollution tolerant populations (ASTM 1996, Lamberson *et al.* 1996).

Like *Leptocheirus, Grandidierella* has a relatively short generation time (30 d @ 19 ºC) and is amenable to laboratory culture. Chronic (28 d) tests incorporating survival, growth and reproductive endpoints have been conducted with this species (Nipper *et al.* 1989, Lamberson *et al.* 1996). However, development of a standardized chronic toxicity test with *Grandidierella* is not currently being pursued.

<u>*Grandidierella japonica* Toxicity Test Strengths, Limitations and Potential Confounding Factors</u>

Grandidierella japonica has been less well tested than many of the other species for which the 10-d protocol has been adapted. *Grandidierella* has been used primarily in regional research studies in southern California (Nipper *et al.* 1989, Bay *et al.* 2001). This species is tolerant of a wide range of grain sizes, and salinities (Table 8). In addition, *G. japonica* is tolerant of unionized ammonia (Table 8), a desirable attribute for sediment testing species. Its tolerance of H_2S is has not been determined. Insufficient data is available to compare this species to the other commonly used species in terms of its relative sensitivity to field collected sediments. Lamberson *et al.* (1996) found that the 10-d test protocol using *G. japonica* was comparable to tests with *L. plumulosus, R. abronius and E. estuarius* in sensitivity to Black Rock Harbor and Pearl Harbor sediments. *Grandidierella* demonstrated comparable sensitivity to fluoranthene to other amphipod species, but was considerably less sensitive than *A. abdita* to copper in water-only exposures. Because it has not been used in larger scale studies where synoptic benthic community analyses have been conducted, the correlative relationship between laboratory survival of *G. japonica* and benthic community structure in field samples had not been investigated. One of the primary strengths of this species is that it is the only west coast amphipod that has been demonstrated to be amenable to long-term laboratory culture, and so has potential for chronic toxicity testing (Nipper *et al.* 1989, Lamberson

et al. 1996). There is, however, no current effort to complete a standardized chronic protocol with this species.

One potential problem with this species is its perceived adaptability to contaminants. Swartz *et al.* (1994) suggested that *G. japonica* was the only amphipod species present in DDT/dieldrin-contaminated sediments in Lauritzen Channel (San Francisco Bay) because the population there had apparently adapted to these pesticides. Anderson *et al.* (2001a) found that *G. japonica* was one of two amphipod species present in Consolidated Slip (Los Angeles Harbor), a highly contaminated toxic hot spot that produced sediments toxic to *R. abronius* and *E. estuarius* in laboratory tests. Other potential factors that may affect this species are listed in Table 8.

Table 6. **Test conditions for conducting a 10-d sediment toxicity test with *Rhepoxynius abronius, Ampelisca abdita, Eohaustorius estuarius, Leptocheirus plumulosus,* or *Grandidierella japonica*.**
(Compiled from U.S. EPA 1994b, ASTM 2000e)

Parameter	Conditions
Test type	Whole sediment, static
Temperature	15ºC: *E. estuarius* and *R. abronius, G. japonica* 20 ºC: *A. abdita* 25 ºC: *L. plumulosus*
Salinity	20‰: *E. estuarius* and *L. plumulosus* 28‰: *A. abdita* and *R. abronius* 34‰: *G. japonica*
Light Quality	Wide-spectrum fluorescent lights
Illuminance	500-1000 lux
Photoperiod	24 h light:0 h dark
Test Chamber	1-L glass beaker or jar w/ ~ 10 cm I.D.
Sediment Volume	175 mL (2 cm)
Overlying water volume	800 mL
Overlying water renewal	None
Size and life stage of amphipods	*A. abdita*: 3 – 5 mm (no mature ♀ or ♂) *E. estuarius* and *R. abronius*: 3 – 5 mm *L. plumulosus*: 2–4 mm (no mature ♀ or ♂) *G. japonica:* 3-6 mm (no mature ♂)
Number of organisms/test chamber	20 per test container
Number of replicates/treatment	Depends on objectives of test; 4 reps. Min.
Feeding	None
Aeration	Yes
Overlying water	Clean natural or reconstituted seawater
Overlying water quality	Temp. daily, pH, NH_3, Salinity, and dissolved oxygen of overlying water at start and end of test. Salinity, NH_3 and pH of porewater.
Test duration	10-d
Endpoints	Survival (reburial optional for *E. estuarius, L. plumulosus,* and *R. abronius*)
Test Acceptability	Minimum mean control response of 90% and satisfaction of performance-based criteria outlined in Table 2.

Table 7. Test acceptability requirements for a 10-d sediment toxicity test with *Rhepoxynius abronius, Ampelisca abdita, Eohaustorius estuarius, Leptocheirus plumulosus,* or *Grandidierella japonica*.
(Compiled from U.S. EPA 1994b, ASTM 2000e)

A. Recommended performance criteria:
   1. Size, life stage, and reproductive stage of amphipods must be within the prescribed species-specific ranges at the end of the test (U.S. EPA 1994b, Section 10.3.4).
   2. Average survival of amphipods in the control sediment must be greater than or equal to 90% at the end of the test.
   3. Salinity, pH, and ammonia in the overlying water and sediment grain size are within tolerance limits of test species.

B. Performance-based criteria for culturing *L. plumulosus* include:
   1. Laboratories should perform monthly 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms. If reference-toxicity tests are not conducted monthly, the lot of organisms used to start a sediment test must be evaluated using a reference toxicant (U.S. EPA 1994b, Section 9.16).
   2. Records must be kept on frequency of restarting cultures.
   3. Laboratories should record the pH and ammonia of the cultures at least quarterly. Dissolved oxygen and salinity should be measured daily. Temperature should be recorded daily.
   4. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.

C. Performance-based criteria for field-collected amphipods:
   1. Laboratories should perform reference-toxicant tests on each batch of field-collected amphipods received for use in sediment tests (U.S. EPA 1994b, Section 9.16).
   2. Acclimation rates to test salinity and temperature should not exceed 3 ºC and 5‰ per 24 h.
   3. Amphipods received from commercial suppliers must exhibit active swimming behavior upon placement in water, have full digestive tracts, and display acceptable color.

D. Additional requirements:
   1. All test organisms must be from the same source.
   2. It is desirable to start a test as soon as possible after collection of sediment from the field.
   3. All test chambers should be identical and should contain the same amount of sediment and overlying water.
   4. Negative—control sediment must be included in a test.
   5. The time-weighted average of daily temperature readings must be within ± 1 ºC of the desired temperature. The instantaneous temperature must always be within $\pm$ 3 ºC of the desired temperature.

Table 7 (continued). Test acceptability requirements for a 10-d sediment toxicity test with *Rhepoxynius abronius, Ampelisca abdita, Eohaustorius estuarius, Leptocheirus plumulosus,* or *Grandidierella japonica*.
(Compiled from U.S. EPA 1994b, ASTM 2000e)

6. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

Table 8. Strengths, limitations, and potential confounding factors associated with 10-d amphipod tests.

Species	Chronic Test	Ecological Relevance*	Laboratory Culture	Confounding Factors	Widespread Use	*TIE USE*	10-d H_20 Culture
Rhepoxynius abronius	**N**	**Y**	**N**	**S, GS, NH_3, H_2S (Table 11)**	*y*	**N**	**N**
Ampelisca abdita	**P**	**Y**	**Y**	**NH_3, H_2S (Table 11)**	**Y**	**Y**	**N**
Eohaustorius estuarius	**N**	**Y**	**N**	**NH_3, H_2S (Table 11)**	**Y**	**Y**	**Y**
Leptocheirus plumulosus	**Y UD**	**Y**	**Y**	**NH_3, H_2S (Table 11) feeding**	**Y**	**N**	**Y**
Grandidierella japonica	**P**	**ND**	**Y**	**NH_3, H_2S, (Table 11) adaptation**	**N**	**N**	**Y**

*****Ecological relevance refers to studies demonstrating correlations between laboratory effects and impacts on benthic community structure.
P = Potential; UD = Under Development; S = Salinity, GS = Grain Size; NH_3 = unionized ammonia; H_2S = Hydrogen Sulfide; N = No; Y = Yes; ND = Not determined
Feeding = test results may be influenced by feeding rates; Adaptation = Possible adaptability to contaminants.

Relative Contaminant Sensitivities of Marine Sediment Toxicity Tests with Amphipods

The relative sensitivities of 10-d toxicity tests using the five amphipod species (*Rhepoxynius abronius, Ampelisca abdita, Eohaustorius estuarius, Leptocheirus plumulosus, Grandidierella japonica*) can be compared by assessing responses in water-only exposures (*e.g.*, reference toxicant tests), and by comparing side-by-side sediment tests. Results of water-only exposures were compiled from those reported in the literature (*e.g.*, ECOTOX database and other sources) and in some cases, from unreported experiments conducted at testing laboratories contacted by UC Davis. Results of these experiments are generally reported as 96-h LC or EC50 values except where noted (referenced in the US EPA Ecotox database posted at: www.epa.gov/ecotox). Note that "sensitivity" data are influenced by a variety of factors including organism stress, toxicant bioavailablity, between replicate variability, etc. Water-only exposure data is considered to be more reflective of organism sensitivity, because there are fewer confounding factors in these experiments. However, the results of these experiments may be also be affected by variations in test salinities, because trace metal bioavailability and organism sensitivity is sometimes salinity-dependant. For example, tests conducted at lower salinities may have higher Cd^{2+}than those conducted at higher salinities. In addition, some apparently "euryhaline" species may be more sensitive to chemicals at lower salinities due to osmotic stress (*e.g.*, Anderson *et al.* 1995). Because of these considerations, the following is considered to be a comparison of method sensitivity, rather than species sensitivity.

Comparison of available water-only exposure data indicates variable chemical sensitivity among the five amphipod species. A considerable database exists for cadmium chloride, because this is the most commonly used reference toxicant. The amphipods *L. plumulosus, A. abdita, R. abroni*us are the most sensitive to this trace metal in water-only exposures (posted at: www.epa.gov/exotox), while *G. japonica* and *E. estuarius* are the least sensitive (ignoring the possible influence of salinity). *Ampelisca abdita* is also relatively sensitive to copper, and to a lesser extent zinc in water-only exposures, although the lack of copper and zinc response data for the other

amphipod species precludes further comparison. Meador (1993) found that *E. estuarius* was considerably more sensitive than *R. abronius* to the metalloid tributyltin (TBT).

More comparative data is available for polyaromatic hydrocarbons (PAHs). *Eohaustorius estuarius* and *L. plumulosus* demonstrated comparable sensitivity to phenanthrene, and all species demonstrated similar sensitivities to fluoranthene (*E. estuarius* results not conclusive). DeWitt *et al.* (1989) found that sensitivity of *E. estuarius* to sediment-spiked fluoranthene was comparable to *R. abronius*. A limited amount of data is available for pesticides. Werner and Nagel (1997) found that *A. abdita* and *R. abronius* had comparable sensitivities to the organophosphorous pesticide diazinon (posted at www.epa.gov/ecotox), and Word *et al.* (1987) found that *R. abronius* was sensitive to DDT in the low µg/L range. This comparison is constrained by the lack of comprehensive dose-response data for these and other pesticides using other amphipod species.

Water-only dose-response data are also available for unionized ammonia and hydrogen sulfide, two common constituents of marine sediment porewaters. As opposed to the anthropogenic contaminants listed above, H_2S and NH_3 are naturally produced compounds that may confound interpretation of sediment toxicity test results. Therefore, species that are tolerant of these compounds but are sensitive to anthropogenic contaminants are preferable, except in cases where H_2S or NH_3 are of specific concern (*e.g.*, where effects of high anthropogenic organic enrichment is of concern). The amphipods *A. abdita* and *L. plumulosus* are more sensitive to unionized ammonia than *E. estuarius* and *G. japonica*, and to a lesser extent *R. abronius*. Thus, *E. estuarius* and *G. japonica* are preferable because they are more tolerant of ammonia than *A. abdita* and *L. plumulosus*. Less data is available for H_2S. Knezovich *et al.* (1996) found that *R. abronius* is considerable more sensitive to this sediment constituent than *E. estuarius*; *E. estuarius* is therefore preferable in this regard.

Because of behavioral and life history differences between species, it is also important to assess relative sensitivity to solid-phase sediments. For example, some amphipod species build mucus-lined tubes or burrows and may not be exposed to porewater contaminants to the same degree as free-burrowing species. In addition, those species that are primarily surface detritivores may not be exposed to the same

contaminant concentrations as predaceous species, or as species that primarily consume subsurface particulate matter. A number of studies have used field-collected sediments contaminated with complex chemical mixtures to make comparisons among amphipod tests. Schlekat *et al.* (1995) conducted round-robin tests with *A. abdita*, *E. estuarius*, and *L. plumulosus* using dilutions of Black Rock Harbor, Connecticut sediment (BRH). Although this study was designed primarily to assess inter-laboratory variability rather than relative sensitivity, the results suggest that the three species demonstrated comparable sensitivities to BRH sediment. In this study there was some variability between laboratories at lower BRH concentrations. At concentrations above 25%, all species classified the BRH sediment as toxic. The LC_{50} values were 16.2, 26.5, and 21.1 percent BRH sediment, for *A. abdita*, *E. estuarius*, and *L. plumulosus*, respectively.

Anderson *et al.* (1997) conducted comparisons between *R. abronius* and *A. abdita* at 25 stations as part of an EMAP study in southern California bays and estuaries. In this study, 12% of the stations tested were significantly toxic to *A. abdita* while 40% were toxic to *R. abronius*. There was concordance between the two species on the presence or absence of toxicity at 72% of the stations.

Some studies have compared the relative sensitivities of *E. estuarius* and *A. abdita* to field–collected sediments. Anderson *et al.* (1999) assessed survival of both species to sediments collected from five stations located along a contamination gradient in Moss Landing Harbor, California. In this study, no significant toxicity was detected by *A. abdita* while low survival of *E. estuarius* occurred at the 3 most contaminated stations. Toxicity Identification Evaluations conducted on sediment from the most contaminated station indicated that a non-polar organic compound was responsible for the observed toxicity. Bulk-phase chemical analysis of this sediment found elevated concentrations of organochlorine pesticides, relative to sediment quality guideline values. Gries (2000) compared the relative sensitivity of *E. estuarius* and *A. abdita* as part of a dredge material assessment study at a U.S. Navy facility in Puget Sound, Washington. This study also found that significant mortality of *E. estuarius* occurred in sediment samples that did not inhibit survival of *A. abdita*. Hunt *et al.* (*in press*) found comparable survival of both species exposed to contaminated sediments from Islais Creek and Castro Cove, two sediment hotspots in San Francisco Bay. Weston (1995) compared the relative

sensitivities of *E. estuarius* (10-d), *A. abdita* (17-d), and *R. abronius* (10-d) to sediments spiked with cadmium, DDT, and crude oil. The ranking of sensitivity of these three species to the three contaminants were as follows (more sensitive > less sensitive): cadmium (*R. abronius>A. abdita>E. estuarius*), DDT (*E. estuarius>A. abdita>R. abronius*), crude oil (*R. abronius = A. abdita = E. estuarius*).

Application of Marine Whole-sediment Amphipod Toxicity Tests

Compiled from ASTM 2000e, U.S. EPA 1994b

Because laboratory toxicity tests are designed to be closed, simplified systems, there are inherent limitations to extrapolation of results from these tests to field assessments (*e.g.*, Luoma and Ho 1993, Luoma 1996). Despite this limitation, toxicity tests are a primary tool used in risk-based decision-making, because they provide a measure of the integrated effects of complex mixtures of contaminants. The following discussion summarizes the information currently available for the various whole-sediment toxicity tests for which standardized protocols have been developed, and provides some guidance on the applicability of the various marine species and protocols.

In a review of the toxicity test procedures available for sediment ecological risk assessments (SERAs), Ingersoll *et al.* (1995a) listed several criteria for selection of measurement endpoints that have the least inherent uncertainty for decision making. Several categories of tests were evaluated: whole sediment exposures with benthic organisms, whole sediment exposures with pelagic organisms, and exposures that use organic extracts, suspended solids, elutriates, and porewater. Evaluation criteria included test precision, ecological relevance, causality (*i.e.*, correctly identifying stressors), sensitivity, interference, standardization, discrimination, bioavailability, and field validation. Tests using different sediment phases were given low to high rankings within each evaluation criterion (or, where appropriate, the lack of knowledge for a specific criterion was noted). These authors determined that whole sediment tests with benthic organisms provided the most ecologically realistic phase for assessing organism response. Toxicity test endpoints were also evaluated, including survival, growth, reproduction, behavior, life tables, development, and biomarkers. They concluded that uncertainty associated with survival is less than that of the other endpoints commonly

used in sediment toxicity tests, because mortality is an extreme response with obvious ecological consequences.

These authors noted their emphasis on endpoints (rather than species) was not intended to imply that any particular endpoint is preferable over another with respect to ecological relevance or overall sensitivity. They suggested that a chronic test with a relatively insensitive species (*e.g.*, growth or reproduction with polychaetes) might be far less ecologically relevant or sensitive compared to an acute survival test using a sensitive species (*e.g.*, survival of amphipods).

Given the extensive information supporting the utility of the 10-d amphipod survival test in sediment quality assessments, and the conclusion by Ingersoll *et al.* (1995a) that the sediment phase and test endpoints in this protocol present analysts with the least amount of uncertainty, this protocol should be included in any SERA. Although the standard 10-d protocol is appropriate for all of the species listed above, there are specific situations where some species may be more appropriate than others for determining potential ecological effects.

There is insufficient comparative data to adequately evaluate the relative contaminant sensitivity of tests with the five amphipod species. Water-only exposure data, and the limited number of field sediment comparisons, suggest that with some sediments, there may be species-specific differences in response. There is also some evidence to suggest that there may be differences in response due to variations in amphipod behavior. Primarily free-burrowing amphipods (*e.g.*, *Eohaustorius estuarius, Rhepoxynius abronius*) may detect toxicity better than tube-building species (*Ampelisca abdita*), particularly when moderately contaminated sediments are tested (*e.g.*, Anderson *et al.* 1997, Anderson *et al.* 2000, Gries 2000). There is less comparative data for those species that construct burrows (*L. plumulosus, G. japonica*), but we expect that free-burrowing species will provide better detection of contaminant effects (or, less uncertainty involving decisions of potential ecological risk). Until more experimental evidence is available, a conservative approach to species selection for sediment assessments is recommended. Assessments of moderately contaminated or uncharacterized sediments should include a free-burrowing species such as *R. abronius* or *E. estuarius*. Both are appropriate for sandy (< 90 percent silt/clay) marine sediments,

and *E. estuarius* is appropriate for silty estuarine sediments having grain sizes and salinities beyond the tolerance range for *R. abronius*. Any of the other species may also be used for sediment assessment, but it is advisable that *A. abdita*, *L. plumulosus*, and *G. japonica* should be tested synoptically with the free-burrowing species as part of initial sediment toxicity assessments to confirm comparable sensitivity to the field sediments under consideration. If *G. japonica* is used, test animals must be collected from a population inhabiting a pristine environment (Lamberson *et al.* 1996). In specific situations where acute toxicity is not expected, the chronic test with *L. plumulosus* may also be appropriate because this is the only species for which a chronic protocol has been thoroughly evaluated. However, longer-term tests with sublethal endpoints do not always provide additional information. Using sediments spiked with a number of compounds, Farrar *et al.* (1999) found that the 10-d test with *L. plumulosus* was equally or more sensitive than the 28-d test with this species. (Note: chronic toxicity of low salinity sediments (< 15‰) may also be tested with the amphipod *Hyalella azteca*; see discussion of freshwater sediment tests).

Because of concerns about the effects of sediment grain size on amphipod survival, study designs using these species, particularly *R. abronius*, should include reference sediments having comparable TOC and particle size distributions as the test sediments. Recent research suggests that rather than characterizing grain size distributions as percent fines, the silt and clay fractions should be quantified because these may better correlate with amphipod mortality (*e.g.*, Hunt *et al.* 2001). In addition, interstitial and overlying water unionized ammonia and hydrogen sulfide should always be measured. Other confounding factors may also affect organism response to sediments and these should be also considered (see Table 3; discussion in Ingersoll 1995a).

Depending on the contaminants of concern, additional toxicologic information regarding exposure routes may be obtained by exposing amphipods to both solid-phase sediment and porewater. For example, Anderson *et al.* (2000) found low survival of *E. estuarius* in whole sediment collected from San Francisco Bay, but no effects in porewater extracted from these samples. These authors concluded that contaminant exposure occurred via particle ingestion rather than dermal uptake. *E. estuarius* can be tested in long-term water-only exposures (*e.g.*, $\geq$ 10 days), and may be more useful for

this type of comparison than species that demonstrate poor survival in water-only exposure (*e.g.*, ≤ 96-h; *R. abronius, A. abdita*). In some cases, additional endpoints and measures may complement amphipod tests. These may include biomarker endpoints (*e.g.*, Werner and Nagel 1997), and measurements of tissue contaminant concentrations in laboratory-exposed amphipods (Anderson *et al.* 2000).

Marine Polychaete Sediment Toxicity Tests

Polychaete (*Nereis (Neanthes) arenaceodentata)* 28-d Growth and Survival Toxicity Test
Compiled from U.S. EPA 1998, ASTM 2000f, Dillon *et al.* 1993

Polychaetes are ecologically important members of benthic infaunal communities and are an important food resource for a number of marine invertebrate, fish, and bird species world-wide. The annelid *Nereis (Neanthes) arenaceodentata* is amenable to laboratory culture and has been used extensively in sediment toxicity studies for a number of years (Reish 1985, Pesch and Hoffman 1983, Dillon *et al.* 1993). The ability to routinely produce known-age juveniles has allowed development of a chronic (20-d or 28-d) sublethal toxicity test protocol with this species that incorporates growth and survival endpoints. Guides for a chronic toxicity test with *N. arenaceodentata* are provided in ASTM (2000f) and U.S. EPA (1998), and these cite the 20-d procedure of Johns *et al.* (1990) and the methods of Dillon *et al.* (1993) as example protocols. Recent research by the Army Corps of Engineers has led to some suggested modifications of the chronic protocol, (*e.g.*, Dillon *et al.* 1993 and 1995, Bridges *et al.* 1997). The following summarizes methods for the 10-d acute procedure and gives a brief description of the 20-d chronic protocol developed by Johns *et al.* (1990). Recent research suggesting modifications of the chronic protocol is also discussed. Because of concerns regarding feeding effects and regulatory implications concerning interpretation of sublethal endpoints, U.S. EPA (1998) do not recommend the chronic test with this species for dredge material testing, except for TIER IV studies where additional information is needed.

Juvenile polychaetes used for marine sediment toxicity assessments are provided by commercial suppliers from laboratory-reared populations. Test containers are usually

one-liter glass beakers or jars containing 2 cm of homogenized sediment and filled with uncontaminated overlying seawater adjusted to the appropriate salinity (28 ‰ ± 2 ‰) and temperature (20 ºC); smaller test containers and sediment volumes have also been used (*e.g.*, Green *et al.* 1999). Test sediment and overlying water are allowed to equilibrate for 24 hours, after which five, 2-3 week-old post-emergent worms are placed in each beaker. Test chambers are aerated gently, and overlying water is generally not renewed in the 10-d test, except in situations where test sediments have unionized ammonia concentrations above protocol thresholds. In these cases it is sometimes necessary to purge ammonia by renewing overlying water before worms are added to the beakers. A negative control sediment consisting of five laboratory replicates of negative control sediment (*e.g.*, amphipod home sediment), is included with each sediment test. Five laboratory replicates of each sample are tested for ten days, after which the samples are sieved and the surviving worms are counted.

A chronic toxicity test with this species is still under development. The 20-d test described by Johns *et al.* (1990) is identical in design to the 10-d protocol described above. To account for initial worm biomass, three replicates of 5 worms each are isolated and dried at the beginning of the test. Sediment overlying water in the 20-d test is renewed every third day; 300 mL test water is removed and replaced with 28‰ aerated seawater. Feeding rates vary between protocols. Johns *et al.* (1990) recommend feeding every-other day with 8 mg Tetramin™ flake food per worm. More recent studies have indicated that feeding rates may influence test sensitivity, particularly in chronic exposures assessing weight as an endpoint (Bridges *et al.* 1997). These authors suggest a reduced, twice-weekly feeding rate using a slurry consisting of 2 mg Tetramin™ flake food and 1 mg alfalfa per worm, per feeding. After 20 days, the sediments are sieved to recover the test animals, and the number of survivors is recorded for each replicate (Dillon *et al.* 1993, and Bridges *et al.* 1997 recommend a 28-d exposure period). At the end of 20 (or 28) days, the sediment is sieved, live animals are counted and placed in distilled water to rinse and separate them from their tubes. The worms are then dried and weighed. In addition to feeding modifications described above, Bridges *et al.* (1997) found that the sensitivity of the chronic 28-d protocol was greater when younger (<48-h)

worms were used, and when the exposure time was increased to 7 weeks. Fifteen replicates were used in these experiments, each containing one worm.

Table 9. **Test conditions for conducting a 10-d sediment toxicity test with *Nereis (Neanthes) arenaceodentata*.**
(Compiled from ASTM 2000.)

Parameter	Conditions
Test type	Whole sediment, static, non-renewal
Temperature	20ºC
Salinity	28-36‰
Light quality	NA
Illuminance	NA
Photoperiod	NA
Test chamber	Glass 1-L beaker with 10-cm internal diameter, covered
Sediment depth	2 cm
Overlying water volume	Fill to 750-mL mark
Overlying water renewal	None
Life stage	Uniform; young adults or 2-3 week emerged juveniles
Number of organisms/test chamber	5
Number of replicates/treatment	5
Feeding	None
Aeration	Yes
Overlying water	Clean natural or diluted (filtered to at least 5 μm) seawater, or reconstituted seawater
Overlying water quality	Temperature and salinity daily; pH, NH_3, and dissolved oxygen of overlying water at start and end of test. pH of porewater at start and end.
Test duration	10-d
Endpoints	Survival
Test Acceptability Criteria	Mean control survival 90% (minimum rep control survival = 80%) and satisfaction of performance-based criteria outlined in Table 10.

Table 10. **Test acceptability requirements for a 10-d sediment toxicity test with *Nereis (Neanthes) arenaceodentata*.**
(Compiled from ASTM 2000.)

---

1. Mortality during pre-test holding period should be < 5%.
2. Test chambers must be identical.
3. Treatments and test organisms must be assigned randomly to test chambers.
4. Required controls must be included.
5. All test organisms must be from the same population.
6. Dissolved oxygen (DO) must be maintained at or above 60% saturation. When aeration is discontinued, DO saturation can drop below this level.
7. Individual temperature readings must not vary by more the 3ºC from test temperature, and time-weighted average should be within 1ºC for duration of test. Concurrent measurements should not vary by more than 2ºC.
8. Wild polychaetes were maintained in the laboratory for more than 2 weeks without demonstration of lack of effect of holding time on sensitivity.
9. Any solvent used affected survival.
10. Analytical methods for measuring concentration of test material were not validated.

Relative Contaminant Sensitivities of Marine Sediment Toxicity Tests with Polychaetes

The relative sensitivity of the *N. arenaceodentata* test may be assessed by comparing water-only exposure data to that of the commonly used marine sediment test species. This test was considerably less sensitive than tests with 4 of the 5 amphipods species to cadmium, and had comparable sensitivity to *Eohaustorius estuarius*. The *N. arenaceodentata* test was more sensitive than the amphipod *G. japonica* to copper, but was less sensitive than *A. abdita* to both copper and zinc. The *N. arenaceodentata* test was also less sensitive to phenanthrene than tests with the amphipods *L. plumulosus* and *A. abdita*, and was considerably less sensitive to fluoranthene than all five amphipod species tested with this compound. The *N. arenaceodentata* test was less sensitive to PCBs (Arochlor 1254) than *A. abdita*. *N. arenaceodentata* was also less sensitive to both hydrogen sulfide and un-ionized ammonia than the amphipods (US EPA,).

A number of studies have compared the relative sensitivity of chronic tests with *N. arenaceodentata* to acute and chronic tests with various amphipod species using field-collected sediments. Using samples from three Superfund sites in Puget Sound, Pastorak and Becker (1990) compared the sensitivity of the 10-d amphipod survival protocol using either *R. abronius* or *E. estuarius* to the *N. arenaceodentata* 20-d survival and growth protocol. Their results demonstrated that the 10-d protocol using either amphipod species was more sensitive than either endpoint using the polychaete protocol. Anderson *et al.* (1998b) tested sediment samples collected throughout California with the 10-d amphipod protocol using *Rhepoxynius abronius* and the 20-d growth and survival protocol using *N. arenaceodentata*. Of the 341 samples tested, 78% significantly inhibited amphipod survival, whereas 2 and 26% significantly inhibited *N. arenaceodentata* survival and biomass, respectively. Amphipod mortality in this study was significantly correlated with a number of bulk-phase contaminants, including metals, pesticides, PAHs, and PCBs. No significant correlations were determined between polychaete survival or biomass and measured contaminants. These authors found that the statistical power of the polychaete protocol, as defined by the 90th percentile Minimum Significant Difference (MSD) detected by this protocol, was considerably lower than that determined for the amphipod protocol. Therefore, although the protocol using *R. abronius* was a more sensitive indicator of toxicity, this was apparently due in part to

greater statistical power, rather than greater sensitivity of the organisms or endpoints. Green *et al.* (1999) assessed the sensitivities of a 28-d growth and survival test using *N. arenaceodentata* to the 28-d growth, mortality, and reproduction protocol using the amphipod *Leptocheirus plumulosus*. These authors found the amphipod to be considerably more sensitive than the polychaete to trinitrotoluene (TNT), particularly when comparing the survival and growth endpoints.

Recent studies have indicated that sensitivity of the chronic toxicity test protocols with *N. arenaceodentata* are influenced by a variety of factors, including worm age, exposure duration, experimental design (number of replicates), and feeding regimes. Bridges *et al.* (1997a) concluded that the sensitivity of the growth endpoint of a chronic protocol with *N. arenaceodentata* to Black Rock Harbor (Connecticut) sediment was comparable to amphipod mortality in 10-d exposures when young (<48-h) worms were exposed for longer durations (7-weeks), and the feeding regimes were reduced (twice per week).

Polychaete Toxicity Test Strengths, Limitations and Potential Confounding Factors

The polychaete *N. arenaceodentata* has been used extensively in sediment toxicity assessments. This species is simple to culture, and emergent juveniles of a known age are readily produced for routine testing. Considerable effort has been devoted to developing a chronic growth and survival protocol using this species. Because this species naturally occurs in a variety of estuarine and marine habitats world-wide, it is more tolerant of some of the sediment features that may affect other sediment toxicity testing species, including high silt/clay sediments, and elevated concentrations of hydrogen sulfide and unionized ammonia (Dillon *et al.* 1993). A 28-d protocol with *N. arenaceodentata* has been subjected to interlaboratory testing (Pesch and Hoffman 1983); however, recent versions of the chronic protocol have not been assessed in interlaboratory testing. This species is useful for bioaccumulation testing because it survives well in contaminated sediments and provides adequate tissue for chemical analysis. This species has also been demonstrated to be useful in genotoxicity and cytotoxicity studies (Pesch 1983, Anderson *et al.* 1990).

Protocols using *N. arenaceodentata* may be limited by lower sensitivity relative to some of the other commonly used protocols, particularly in studies of the effects of moderately toxic sediments. The ACOE/EPA inland dredge testing manual (U.S. EPA 1998) does not consider the 10-d test with this species to be a sensitive "benchmark" test, and recommend that for assessing toxicity of dredge materials, this protocol only be used in conjunction with other benchmark protocols (*e.g.*, amphipods). Although adjustments to the chronic protocol for *N. arenaceodentata* suggest it may be as sensitive as some of the 10-d amphipod tests when designed properly (Bridges and Farrar 1997), more research is needed to evaluate whether additional toxicologic information is provided with this protocol relative to information from acute tests with amphipods. Dillon *et al.* (1993) demonstrated the potential ecological significance of the growth endpoint with *N. arenaceodentata* by determining the minimal daily growth necessary for reproduction to occur with this species. However, because the chronic protocol has not been used in larger scale studies where synoptic benthic community analyses have been conducted, the correlative relationship between laboratory survival and growth of *N. arenaceodentata* and benthic community structure in field samples has not yet been determined.

Application of Marine Polychaete Whole-sediment Toxicity Tests

Compiled from ASTM 2000f, Johns *et al.* 1990, Dillon *et al.* 1993

Relative to criteria used by Ingersoll *et al.* (1995a; *i.e.*, a solid phase test with a burrowing species), acute and chronic toxicity test protocols with *Nereis (Neanthes) arenaceodentata* provide less uncertainty regarding risk than tests using other sediment matrices and non-infaunal species. However, the 10-d survival protocol using this species is probably not sufficiently sensitive (see U.S. EPA 1998). Recent research (Bridges and Farrar 1997) has suggested longer-term tests (> 20-d) with this species may be more sensitive when appropriate design parameters are adjusted (*i.e.*, exposure duration, worm age, replication, feeding regime), and when growth is used as an endpoint. However, no revised standardized chronic protocol is available, and there is insufficient information to determine whether a revised chronic polychaete protocol would provide additional toxicologic information relative to the other solid-phase protocols. As discussed above for the amphipod protocols, if a chronic protocol using

this species is used in a sediment ecological risk assessment, a 10-d amphipod tests using one of the free-burrowing species should also be included. Because *N. arenaceodentata* is tolerant of H_2S and NH_3, it may be useful in assessments of sediments where these compounds are present at sufficient concentrations to confound interpretation of amphipod test results. In addition, chronic tests with this species may also be useful if genotoxic, cytotoxic or bioaccumulative chemicals are of concern.

MARINE SEDIMENT POREWATER TESTS

Water between particles in sediment is called porewater water. Marine sediment toxicity tests using porewater are all short term tests and are reviewed as follows:

Short-term Embryo-larval Toxicity Tests: Purple Sea Urchin (*Strongylocentrotus purpuratus*) and Bay Mussel (*Mytilus galloprovincialis*)

Depending on the organisms and contaminants present, porewater is believed to be an important route of exposure for many species and life stages, particularly in sediments contaminated by more water-soluble (*i.e.*, low K_{ow}) chemicals. Porewater tests allow the use of a variety of protocols developed for aqueous samples. A number of species and protocols used in water-only toxicity assessments have been used to test marine porewater toxicity (see Carr *et al.* 2003, and references therein). Three of the standardized water-only protocols described above are most often used for porewater toxicity assessments. These are the purple sea urchin (*S. purpuratus*) fertilization test, and the embryo-larval development tests with either the purple sea urchin or bivalve mollusk (*M. galloprovincialis*). A brief description of methods for using these protocols in sediment porewater testing is provided here.

Porewater is extracted from sediment by either centrifugation or various pneumatic techniques. Once extracted, it is refrigerated, or in some cases frozen, and held for testing. Potential confounding factors associated with porewater manipulations are discussed later. Because it often requires relatively large volumes of sediment to provide sufficient volumes for porewater testing, these tests are usually conducted in smaller containers such as scintillation vials, using smaller volumes of porewater (*e.g.*, $\leq$ 20 mL per replicate).

Tests with bivalves and sea urchins are conducted as described previously. Because gametes and embryos of invertebrates are sensitive to a number of physico-chemical factors associated with porewater, it is sometimes necessary to adjust porewater prior to addition of animals. Low salinity samples are adjusted to the appropriate salinity with the addition of hypersaline brine or artificial sea salt. Porewater with pH beyond the species tolerance range is sometimes adjusted with the addition of HCl or NaOH. Various dilutions of porewater are often tested (*e.g.*, 100%, 50%, and 25%), so that a dose-response relationship can be determined. Porewater is diluted with control (uncontaminated) seawater. Testing at lower porewater concentrations may also minimize the effects of confounding factors such as unionized ammonia, hydrogen sulfide, and extremes in pH.

Relative Contaminant Sensitivities of Marine Sediment Porewater Toxicity Tests

The relative sensitivity of the purple sea urchin echinoderm fertilization and development test and the bivalve mollusk embryo development test in water-only exposures was discussed in the section on marine water column tests. The following discussion gives examples of studies where the relative sensitivity of these tests were compared to each other and to other test protocols used in marine sediment toxicity assessments.

Anderson *et al.* (1997) assessed sediment quality in selected Southern California bays and estuaries using two amphipods (*Rhepoxynius abronius* and *Ampelisca abdita*) exposed to whole-sediment, and the purple sea urchin (*Strongylocentrotus purpuratus*) fertilization and embryo development tests using porewater. Porewater toxicity was compared using concentration of 100%, 50%, and 25% porewater. Estimates of the spatial extent of toxicity in this study were based on Cumulative Distribution Functions; 58% of the area sampled significantly inhibited survival of the amphipod *R. abronius*, and 11% of the area sampled significantly inhibited survival of *A. abdita*. The embryo-larval test was considerably more sensitive than the amphipod tests. At 100%, 50%, and 25% porewater concentrations, the percent area significantly toxic to purple sea urchin embryo-larval development was 91%, 83%, and 51%, respectively. The sea

urchin fertilization test was less sensitive than the embryo development test. Using 100% porewater, 43% of the area sampled was significantly toxic to sea urchin sperm.

Fairey *et al.* (1998) conducted sediment quality assessments in San Diego Harbor, California, using the amphipod *R. abronius*, and sea urchin (*S. purpuratus*) embryo development in porewater. Amphipod survival was significantly inhibited in 57% of the samples tested in this study; sea urchin development in 100% porewater was significantly inhibited in 74% of the samples tested. As expected, the toxicity test protocols used in this study apparently responded to different chemical constituents in these samples. Of the 164 samples tested with both protocols, concordance between the two protocols on the presence or lack of toxicity was achieved in 30% of the samples.

Anderson *et al.* (2001a) assessed toxicity of Los Angeles Harbor whole-sediment samples using amphipods (*R. abronius* or *E. estuarius*), and porewater toxicity using embryos of the red abalone (*Haliotis rufescens*). The 48-h abalone embryo-larval development test is analogous to the *M. galloprovincialis* embryo-larval development test, and demonstrates comparable sensitivity (Hunt and Anderson 1993). While 29% of the sediment samples were toxic to amphipods, 79% were toxic to abalone embryos exposed to 100% porewater. Many of the porewater samples in this study had un-ionized ammonia concentrations above the effect threshold for the abalone development test. To minimize the influence of ammonia, porewater was also tested at 50% and 25% concentrations. When tested with 50% porewater, 40% of the samples were toxic, and when tested with 25% porewater, 14% of the samples were toxic.

Long and Buchman (1989) compared the relative sensitivity of embryo-larval development tests using *S. purpuratus* and *M. galloprovincialis* to sediment elutriates from San Francisco Bay. These authors found that overall, the mussel embryos appeared to be more sensitive, but that there was little concordance between these protocols, suggesting that sensitivity differences depended on the sediment contaminants present**.** In a similar study, Phillips *et al.* (2000) compared the relative sensitivity of these protocols using sediment elutriates from a station in north San Francisco Bay (Grizzly Bay). This study showed that sediment elutriates from Grizzly Bay significantly inhibited mussel larval development, but had no effect on sea urchin larval development.

Toxicity Identification Evaluations showed that sediment toxicity at this station was caused by divalent cations (*e.g.*, copper).

Bay *et al.* (2003) compiled data from a number of Canadian, national and regional studies to compare the relative performance of solid-phase and porewater toxicity tests. A majority of the data used in this comparison represented amphipod (*A. abdita, R. abronius, E. estuarius*) acute toxicity tests, and sea urchin (*Arabacia punctulata, S. purpuratus*) fertilization and embryo development tests in porewater. These authors found that for marine porewater and solid-phase tests conducted side-by-side, the same classification (either toxic or non-toxic) was obtained for 54% of the samples tested. Most of the agreement between tests was due to absence of toxicity in the samples. When only toxic samples were examined, porewater toxicity tests were much more likely than solid-phase tests to detect toxicity. Both test methods identified the same sample as toxic less than 15% of the time. These authors also found that marine porewater toxicity tests provided most of the unique toxicity data, regardless of the protocol or species used. In most studies, the solid-phase test did not detect any toxic samples that were not also identified as toxic on the basis of the porewater toxicity test results.

Marine Sediment Porewater Toxicity Test Strengths, Limitations and Potential Confounding Factors

Protocols for assessing toxicity of marine sediment porewater have been adapted from water-only toxicity test protocols. In California, pore-water test protocols have been used in conjunction with solid-phase tests in regional sediment quality surveys, and more recently, as part of national sediment quality assessment programs such as U.S. EPA's Environmental Monitoring and Assessment Program (EMAP), and NOAA's National Status and Trends program. There are several advantages to using porewater tests to assess sediment quality. For some species and life stages, porewater represents the primary route of exposure (see Carr *et al.* 2003). Porewater tests allow the use of test protocols that use more sensitive species and life stages. Many of these tests are short-term (< 96-h) and therefore provide rapid responses, an attribute that is particularly useful in sediment toxicity screening studies. Because these tests are conducted using an aqueous phase, they are more amenable to established Toxicity Identification Evaluation

procedures. Chemical analyses of porewater allows for comparisons to water quality standards and criteria. Additional practical advantages of porewater testing are discussed in Winger *et al.* (2003).

There are a number of disadvantages to porewater toxicity testing. Porewaters in nature are usually anoxic, but may be oxidized during sample manipulation and toxicity testing. The influence of possible artifacts introduced by sample collection, extraction, and preparation for testing remains poorly understood. An important limitation of porewater as a test matrix is that it excludes consideration of other routes of exposure, such as sediment ingestion (*e.g.*, Forbes *et al.* 1998, Lee *et al.* 2000). The ecological relevance of porewater tests for predicting effects on benthic infauna are questionable because porewaters are often tested with pelagic or epibenthic species and life stages. In some cases, these species and life stages are sensitive to porewater characteristics such as extremes in pH, and elevated concentrations of unionized ammonia and hydrogen sulfide (Bay *et al.* 2003). Other potential confounding factors associated with porewater testing are listed in Table 11. The application of porewater tests has been limited by the lack of standardized protocols for collection, manipulation, extraction, and toxicity testing.

Application of Marine Sediment Porewater Toxicity Tests

Relative to the criteria of Ingersoll *et al.* (1995a) discussed above, porewater tests using early life stages of water column and epibenthic marine invertebrates provide greater uncertainty regarding ecological risk than tests using whole sediment and infaunal species. Despite this, porewater testing may provide useful toxicologic information, particularly when used in studies with additional solid-phase tests and in conjunction with appropriate physico-chemical analyses. While we recognize that a number of species and protocols have been reported for porewater toxicity testing, this summary is limited to tests with standardized embryo-larval toxicity test protocols (U.S. EPA 1995a), particularly those involving the gametes of sea urchins, bivalve mollusks or analogous species, because these are the protocols most commonly used in sediment assessments using west coast marine species.

As discussed above, analyses of statistical correlations provide one way to investigate the relationship between laboratory sediment toxicity test results and potential

impacts on benthic community structure. Correlations between porewater or solid-phase test results with benthic community characterizations have been demonstrated in a number of recent studies. In a survey of sediment toxicity in Puget Sound, Washington, Long *et al.* (1999) found no toxicity using amphipods (*Ampelisca abdita*), but did find inhibition of fertilization (*Arabacia punctulata*) using porewater extracted from a number of samples. In this study, multivariate statistical analyses demonstrated negative correlations between the presence of echinoid species in benthic community samples and inhibition of fertilization in laboratory porewater tests. Anderson *et al.* (2001a) found similar negative correlations between inhibition of mollusk (*H. rufescens*) embryo development in the laboratory, and the number of molluscan species and individuals in Los Angeles Harbor benthos. Carr *et al.* (2000) also found that inhibition of sea urchin embryos (*A. punctulata*) in porewater was negatively correlated with benthic community metrics in samples where amphipod mortality was minimal. These comparisons demonstrate that porewater toxicity tests may provide ecologically relevant information, and that sometimes this information is unique to the porewater test protocol. This may be particularly useful in situations where contamination effects on solid-phase species are not observed.

In a comparison between porewater and solid-phase test protocols, Winger *et al.* (2003) discussed the advantages and disadvantages of porewater toxicity test protocols, and listed a number of recommendations to improve interpretation of test results. These authors suggested that porewater tests be used in conjunction with solid-phase tests when possible, to provide a weight-of-evidence approach to sediment quality assessment. Because of the potential influence of confounding factors such as elevated NH_3 and H_2S, and extremes in pH, these authors suggest that these and other parameters (Table 8) need to be routinely quantified. More research is needed to increase our understanding of the influence of non-contaminant factors on porewater test response. This is particularly important when unexplained toxicity occurs at reference sites. In addition, rather than relying on bulk-chemistry methods, chemical analyses of porewater should be conducted in conjunction with porewater toxicity tests because this provides a better direct measure of exposure for porewater test species. This should include measurement of additional binding phases (*e.g.*, DOC) that may influence chemical bioavailability in porewater.

Porewater chemical analyses will also facilitate interpretation of TIE studies, and may be particularly useful when toxicity may be due to non-contaminant factors, or when the cause of toxicity is not obvious. These authors acknowledge that porewater testing may be less appropriate in situations where the primary chemicals of concern are likely to be tightly sorbed to particles (high K_{ow} compounds), or where the bioavailability of dissolved chemicals is reduced by DOC, organic ligands, or other binding phases. In these cases, contaminant exposure may occur via sediment ingestion, and effects may be underestimated by porewater toxicity tests.

Sediment-water interface (SWI) exposures offer one alternative to porewater tests, particularly in organically enriched sediments where there is concern over the confounding effects of elevated concentrations of ammonia or hydrogen sulfide. The SWI system is also appropriate for sediment ecological risk assessments that are concerned with the impacts of sediment-fluxed contaminants on epibenthic and water column species. As discussed earlier, use of invertebrate embryo-larval development in SWI exposures may provide a more ecologically relevant exposure regime for the assessments of impacts to early life stages of infaunal, epibenthic, and water column species. The use of intact sediment core samples reduces artifacts due to sample manipulation. When resources allow, study designs that incorporate SWI and porewater tests using invertebrate gametes and embryos, combined with information from solid-phase tests (*e.g.*, using amphipods) will provide the least uncertainty regarding ecological risk to marine and estuarine environments.

MARINE SEDIMENT-WATER INTERFACE TOXICITY TESTS

To address some of the disadvantages of porewater toxicity testing, Anderson *et al.* (1996, 2001b) adapted protocols that assay embryonic development to the measurement of sediment toxicity at the sediment-water interface. This test is conducted using the 48-h bivalve embryonic development test or the 96-h sea urchin development test, as described above. In this test, intact sediment cores are collected, and the top of the core is filled with clean overlying seawater (300 mL). A screen tube is then inserted into the core so that the screen sits on top of the sediment. At the initiation of the test, embryos are inoculated into the screen tube. The negatively-buoyant embryos settle to

the bottom and develop on the screen at the sediment- interface. This test is designed to assess toxicity of chemicals fluxed from the sediment into the overlying water. It minimizes the influence of pH extremes, or elevated NH_3 or H_2S, while still using a sensitive early life stage toxicity test that incorporates sublethal endpoints.

Marine Sediment-water Interface Toxicity Test Strengths, Limitations and Potential Confounding Factors

This test system uses the same invertebrate embryo-larval toxicity test protocols as described above for porewater testing, but the animals are exposed to solid-phase sediment samples at the sediment-water interface. The sediment-water interface (SWI) is an ecologically important habitat where significant densities of benthic and epibenthic species occur. In addition to being a likely contaminant exposure location for strictly benthic species, the gametes and embryonic stages of many infaunal, epibenthic, and water column species may spend critical phases of their early development associated with this environment. There are a number of advantages to this exposure system. One is that it may be used to assess toxicity of intact, unhomogenized sediment cores. This minimizes artifacts that may result from the manipulation of sediment and porewater samples. For example, Anderson *et al.* (2001b) used SWI exposures to show that toxicity of homogenized sediments was significantly less than intact (unhomogenized) sediments, and this coincided with lower flux of trace metals from homogenized sediments relative to intact samples. Exposure of invertebrate embryos at the SWI also minimizes effects of NH_3 and H_2S, and eliminates problems with pH extremes that may occur in porewater exposures. Hunt *et al.* (in press) found that SWI tests with purple sea urchin embryos were useful for separating toxic effects of ammonia from those due to anthropogenic contaminants in sediment samples from San Francisco Bay.

One disadvantage of the SWI exposure system is that contaminants fluxing from sediments may be diluted by the overlying water. This system therefore may be underprotective for strictly infaunal species. In addition, artifacts due to the exposure system itself have not been thoroughly investigated (*e.g.*, chemical binding on the screen tube surfaces). This test is relatively new, so the ecological relationship between SWI

toxicity in laboratory exposures and effects on benthic communities has not been determined.

MARINE SEDIMENT BIOACCUMULATION TESTS

Sediment Bioaccumulation Tests with Macoma balthica, Macoma nasuta, Nereis virens, Nereis diversicolor, and Yoldia limatula

The accumulation of chemicals in sediments reduces their direct bioavailability to pelagic organisms but increases exposure of epibenthic and benthic infaunal species. Bioaccumulation tests with sediment species are designed to generate quantitative estimates of steady-state tissue residues. These data are useful in human health and ecological risk assessments, particularly when the chemicals of concern are heavy metals or organic compounds with high octanol-water partitioning coefficients (K_{ow}). This information is useful to estimate bioavailability of sediment-associated contaminants to benthic species in order to assess their direct effects, and to assess potential for contaminant transfer to higher trophic levels. Although a number of marine species have been used for this type of testing, those listed above are the ones most commonly used, and are recommended by ASTM (2000d).

Bioaccumulation tests with sediment are usually conducted for 28 days to estimate steady-state tissue residues, although longer-term exposure methods may be used if 80% of steady-state is not achieved within this time. Sediments tested may be either field-collected or spiked samples. Tests are conducted in aquaria or borosilicate glass food trays supplied with flow-through overlying water. Sediment serves as the habitat and source of food and contaminants for the test organisms. Sediment volumes in each exposure chamber depend on test species requirements. At least 50 g of wet sediment for each 1 g of wet flesh tissue is added initially for surface deposit-feeding bivalves and larger marine deposit feeders. The initial depth for deposit feeding clams such as *Macoma* is 2 to 5 cm. Additional sediment is added to tests with species such as *Macoma* to replenish the bioavailable fraction of sediment consumed by the clams during the exposure. No supplemental food is added to the test containers. All sediments are characterized for contaminant concentrations, TOC, percent sand, silt, clay, and moisture

content, at a minimum. At the end of the exposure period, guts of animals are purged by placing them in clean control sediment, or in some cases clean seawater.

Bioaccumulation in animals exposed to contaminated sediments is compared to that of animals exposed to control and reference sediments. A control sediment contains no or very low contaminant concentrations, and comparison of bioaccumulation in the control sediment provides information on whether contamination from the overlying water or exposure system has occurred. A reference sediment collected from the same region as the test sediment is also compared, and tests with this sediment may be used as an indicator of localized sediment conditions exclusive of the specific contaminants of concern.

The primary criterion for selection of appropriate test species is that they must ingest sediment (infaunal deposit feeders are preferred over epifaunal deposit feeders) and be sufficiently tolerant of contaminants to survive in a 28-d exposure. In addition, test organisms should be large enough to provide sufficient tissue for chemical analysis, and should be of uniform size from similar age classes. A list of candidate species is provided in ASTM (2000d). Tests should not be conducted with gravid individuals.

The experimental objective of bioaccumulation tests is to quantify contaminant bioaccumulation by animals exposed to sediments and determine whether this accumulation is statistically greater than control or reference sediments. A minimum of 8 replicates (exposure chambers) is recommended for this procedure. The simplest design compares test and control sediment results in 24 tissue samples: 8 controls at t_0 (time 0), 8 controls at t_{28}, and 8 test samples at t_{28}. Depending on the size of the test animals, it is sometimes necessary to composite animals within each replicate to provide sufficient tissue for chemical analysis. Lipid levels are characterized in control and test animals for normalization purposes. Examples of percent of steady-state tissue residues for selected neutral organic and metal compounds after 10 and 28-d exposures of a variety of test species are provided in ASTM (2000d).

Marine Bioconcentration/Bioaccumulation Test Strengths, Limitations and Potential

If designed correctly, bioconcentration tests are useful for predicting concentrations of chemicals likely to occur in aquatic organisms in field situations, and

this information is useful for assessing hazard to higher trophic level consumer species. Sediment bioaccumulation tests are useful for determining bioavailability of sediment-associated contaminants, and may also be used to assess hazard to higher trophic level organisms. This information is particularly useful for linking risk of contaminated sediments to epibenthic and water column species.

Results of bioconcentration experiments underestimate uptake via food. This uptake route may be a more important source of residues in fish than water for stable neutral organic chemicals that have a log k_{ow} between 4 and 6 (ASTM 2000c). In addition, bioavailability of chemicals in bioconcentration tests may differ from those in ambient systems because particulate matter is deliberately minimized in these tests.

Bioconcentration and sediment bioaccumulation tests may be affected by the age, physiological condition, sexual maturity, and reproductive condition of the test animals. In addition, sediment bioaccumulation tests are subject to many of the same interferences and confounding factors described for sediment toxicity tests. For example, because these are laboratory experiments, artifacts associated with sediment collection, transport, and processing may affect results of bioaccumulation experiments relative to *in situ* exposures. Chemical solubility, partitioning coefficients, and other physical and chemical characteristics will differ for sediments tested at temperatures other than those at the sample collection site. In addition, changes in ratios between sediments and overlying waters may influence partitioning and accumulation behavior of compounds.

Table 11. Examples of physico-chemical factors potentially influencing porewater and solid-phase toxicity tests.

Effects apply to both freshwater and marine tests, except where noted. (From Winger *et al.* in press.)

Factors	Nature of Effect(s)	Reference
Ammonia	Ammonia toxicity masks toxicity of other chemicals	Schubauer-Berigan and Ankley 1991, Fairey *et al.* 1998
Alkalinity	Toxicity masks toxicity of other chemicals	Lasier *et al.* 1997
Salinity/ Conductivity	Affects bioavailability, changes toxicity Stress	Flegal *et al.* 1994, O'Reilly Wiese *et al.* 1997
Hydrogen sulfide	H_2S toxicity masks toxicity of other chemicals Decreases bioavailability of cations	Leonard *et al.* 1998, Ankley *et al.* 1996, Wang and Chapman 1999
Dissolved organic carbon	Reduces bioavailability	Van Ginneken *et al.* 1999, Green *et al.* 1993
pH	Changes speciation, bioavailability, toxicity (metals)	Bay *et al.* 1993
Alkalinity	Ion toxicity (stress)	Lasier *et al.* 1997
Ferric/ manganese oxides	Affects metal bioavailability	Bufflap and Allen 1995
Particle size distribution	Maybe stressful when different from natural amphipod habitat; influences bioavailability	DeWitt *et al.* 1989
Biological Factors	Predation/indigenous organisms may kill test Organisms; influence contaminant flux	Reynoldson *et al.* 1994
Species sensitivity	Toxicity and exposure	Bay *et al.* 2003
Avoidance behavior	Influences organism exposure to contaminants	Wang and Chapman 1999, Oakden *et al.* 1984
Route of exposure (*e.g.*, ingestion)	Test species differ in contaminant exposure via ingestion and dermal uptake	Lee *et al.* 2000
Burrowing or tube building behavior	May limit chemical exposure	Aller *et al.* 1988
Exposure (test) durations	Differences affect toxicity	Sibley *et al.* 1997a

PART II. FRESHWATER TOXICITY TEST METHODS

FRESHWATER WATER COLUMN TOXICITY TESTS

Freshwater Algae Water Toxicity Tests

Raphidocelis subcapitata (= Selenastrum capricornutum) 96-h Toxicity Test

Compiled from U.S. EPA 1994c

The green alga *Raphidocelis subcapitata* is a microscopic, unicellular, freshwater plant. A starter culture, obtained from a biological supply company, is inoculated into a batch of culture medium and grown to the appropriate age and cell density. The test inoculum is prepared from this stock culture 2 to 3 hours prior to test initiation.

The *Raphidocelis* test is conducted in 125-mL or 250-mL flasks filled with 25 to 100 mL of test solution (depending on the method of mixing employed during the test). Test solutions can consist of filtered freshwater samples, filtered or synthetic freshwater controls, and reference toxicant solutions. The static, non-renewed test proceeds for 96 hours at 25ºC, after which time the growth endpoint is determined by measurement of cell number, chlorophyll fluorescence, light absorbance, or biomass.

Table 12. Test conditions for conducting a 96-h water column toxicity test for growth: *Raphidocelis subcapitata*.
(Compiled from U.S. EPA 1994c)

Parameter	Conditions
Test type	Water column, static, non-renewal
Temperature	25°C
Light quality	"Cool white" fluorescent lighting
Light intensity	86 $\mu E/m^2/s$
Photoperiod	Continuous illumination
Test chamber	125-mL or 250-mL flask
Test solution volume	50 mL or 100 mL (25 or 50 if manually shaken)
Age of test organisms	4 to 7 days
Initial cell density in test chambers	10,000/mL test solution
Number of replicates/treatment	Minimum of 3
Dilution/control water	Algal stock culture medium, enriched uncontaminated source of natural water, synthetic water, or diluted mineral water
Water quality	Temp., pH daily; conductivity, alkalinity, hardness at start; NH_3, Cl if needed
Shaking rate	100 cpm continuous, or twice daily by hand
Test duration	96 hours
Endpoints	Growth (cell counts, chlorophyll fluorescence, light absorbance, biomass)
Test Acceptability Criteria—Negative Control Performance	1×10^6 cells/mL with EDTA, 2×10^5 cells without; variability among reps $\leq$ 20%

Raphidocelis subcapitata Toxicity Test Strengths, Limitations and Potential Confounding Factors

Because EPA recommends screening samples for toxicity with a fish, invertebrate and alga, and because the protocol for *R. subcapitata* is the only freshwater algal protocol, this test has also been used extensively in ambient monitoring. This test is particularly relevant for testing in situations where herbicides may be the primary chemicals of concern (de Vlaming *et al.* 2000), and in cases where eutrophication may be of concern. This test has also been demonstrated to be sensitive to metal-contaminated ambient samples, particularly those associated with mine drainages (de Vlaming *et al.* 2000). These authors reported that results of tests with this alga may be confounded by poor growth rates in control treatments. Because the test with *R. subcapitata* was originally designed to detect toxicity and eutrophication, nutrient concentrations in the control media provide for sub-maximal growth rates. Algal growth is usually higher in ambient samples due to enhancement by naturally occurring nutrients. These authors suggest that poor algal growth in ambient samples may be due to low hardness, alkalinity, or nutrient concentrations, in addition to phytotoxic compounds. In addition, these authors suggest that where toxicants are present, the beneficial effects of the nutrient load may mask growth inhibition of chemicals in ambient samples.

Freshwater Invertebrate Water Toxicity Tests

Rotifer (*Brachionus calyciflorus*) 24-h Embryo Survival Toxicity Test

Compiled from ASTM 2000g

Brachionus is a tiny, free-swimming invertebrate found primarily in freshwater systems. It is an important filter-feeding grazer of phytoplankton, and also serves as a food source for many fishes and invertebrates. Rotifers are valued as a toxicity test organism because of their short, rapid life cycle, sensitivity to contaminants, and availability.

Brachionus embryos used in toxicity testing are hatched from cysts that form as the result of arrested embryonic development. Cysts are hatched by incubation in dilution water at 25ºC for about one day, and produce only female embryos. Cysts can be

obtained from cultured wild populations, but commercially reared strains are recommended, since their sensitivity is well characterized.

The 24-h *B. calyciflorus* test is conducted in darkness at 25ºC in small chambers such as tissue culture plate wells. Each well is filled with 1 mL of test solution and 10 newly hatched neonates. Test solutions can consist of freshwater samples, freshwater controls, and reference toxicant solutions. The test endpoint is determined by counting living and dead rotifers in each test chamber with a dissecting microscope, and determining percent survival.

Brachionus calyciflorus Toxicity Test Strengths, Limitations and Potential Confounding Factors

Because rotifers are ubiquitous in freshwater ecosystems, and play an important ecological role as consumers of phytoplankton and bacteria, they are ecologically relevant pollution indicators. Toxicity tests with *B. calyciflorus* are particularly attractive for benchtop toxicity testing because of their rapid reproduction and short generation times. In addition, rotifer cysts are easily obtained from commercial suppliers and are readily hatched for testing. One limitation of the protocol designed for *B. calyciflorus* is that it is considerably less sensitive to a wide range of chemicals when compared to the other invertebrate toxicity tests. Results with this protocol may therefore not represent risk to more sensitive species.

Table 13. **Test conditions for conducting a 24-h water column toxicity test for embryonic rotifer survival: *Brachionus calyciflorus*.**
(Compiled from ASTM 2000g)

Parameter	Conditions
Test type	Water column, static, non-renewal
Temperature	25°C
Light quality	Darkness
Test chamber	Covered 2.5-mL tissue culture well, or others
Test solution volume	1 mL
Neonate density	10/mL test solution
Neonate age	< 2 hours
Number of replicates/treatment	3
Dilution/control water	Reconstituted freshwater
Water quality	Temperature, pH, hardness at start and end, and dissolved oxygen at start of test
Feeding	None
Aeration	None
Test duration	24 hours
Endpoint	Survival
Test Acceptability Criteria	Negative control survival: ≥ 90% Test chambers identical Organisms randomly distributed to chambers

Cladoceran (*Ceriodaphnia dubia* and *Daphnia magna*) Acute and Chronic Neonate Toxicity Tests

Compiled from ASTM 2000h, U.S. EPA 1993a, U.S. EPA 1994c

Water fleas (suborder Cladocera) are tiny, filter-feeding, freshwater invertebrates inhabiting lakes, ponds and streams world wide. The species *Ceriodaphnia dubia* and *Daphnia magna* are commonly used in water column toxicity testing, and can also be used to test the toxicity of sediments.

Water fleas can be cultured in-house, or gravid animals obtained from commercial suppliers. Short-term tests (24, 48 or 96 hours in duration) are conducted with *C. dubia* or *D. magna*, in 30-mL beakers with 15 mL (*C. dubia*) or 25 mL (*D. magna*) of test solution. A similar 7-d test can also be conducted with *C. dubia*, or in sediment with either species. The 21-day life-cycle test with *D. magna* is performed in larger vessels containing 30 to 40 mL of test solution. Test solutions can consist of freshwater samples, freshwater controls, and reference toxicant solutions. Depending on the test, 1 to 5 neonates less than 24 hours old or 5 days old are introduced into each test chamber, and tests are run at 20 or 25ºC under static renewal or flow-through conditions. Test organisms are fed prior to and during the test with a variety of food items, including yeast + cereal leaves + trout chow (YCT) and unicellular algae (such as *Raphidocelis*). Test endpoints can include survival, growth and reproduction.

Table 14. Test conditions for conducting a short-term (24-, 48- or 96-h) or 7-d water column toxicity test: *Ceriodaphnia dubia* and *Daphnia magna*.
(Compiled from U.S. EPA 1993a, 1994c)

Parameter	Conditions
Test type	Short-term: water column; static renewal or flow-through 7-d *C. dubia*: water column; static renewal
Temperature	20 or 25°C
Light quality	Ambient laboratory illumination
Light intensity	10 to 20 $\mu E/m^2/s$
Photoperiod	16 h light:8 h dark
Test chamber	30-mL covered beaker
Test solution volume	*C. dubia*: 15 mL *D. magna*: 25 mL
Neonate density	Short-term: 5 per beaker 7-d *C. dubia*: 1 per beaker
Neonate age	< 24 hours
Number of replicates/treatment	Short-term: 4 7-d *C. dubia*: 10
Renewal	Short-term: 48 hours 7-d *C. dubia*: daily
Dilution/control water	Synthetic water, diluted mineral water, or uncontaminated natural water
Water quality	Temperature, dissolved oxygen and pH daily; conductivity, alkalinity and hardness at start of test and at renewal; ammonia and chlorine if needed
Cleaning	7-d *C. dubia*: change beakers daily
Feeding	0.1 mL each YCT and algal suspension; daily for 7-d *C. dubia*, pre-renewal for short-term test
Aeration	None
Test duration	Short-term: 24, 48, or 96 hours 7-d *C. dubia*: until 60% of surviving control organisms have three broods (8 days maximum)
Endpoint	Short-term: survival 7-d *C. dubia*: survival and reproduction
Test Acceptability Criteria	Short-term: 90% control survival 7-d *C. dubia*: 80% control survival, with 60% of surviving control adults having had at least 3 brood of 15+ offspring each

Table 15. Test acceptability requirements for a 7-d sediment toxicity test with *Ceriodaphnia dubia* or *Daphnia magna*.
(Compiled from U.S. EPA 1993a, 1994c)

1. The age of test organisms at the start of the test must be within the required range.
2. Hardness, alkalinity, pH and ammonia of overlying water within a treatment should not vary by > 50% during the test.
3. Laboratories should perform monthly 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms. If reference-toxicity tests are not conducted monthly, the lot of organisms used to start a sediment test must be evaluated using a reference toxicant. The duration of the test should be 48 h. Acceptable control survival should be 90%.
4. Laboratories should keep a record of survival of brood organisms and average brood size for each culture and record this information using control charts. Records should also be kept on the frequency of restarting cultures.
5. Laboratories should record the following water quality characteristics of the cultures at least quarterly and the day before the start of a sediment test: pH, hardness, alkalinity, and ammonia. Dissolved oxygen should be measured weekly. Temperature should be recorded daily.
6. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.
7. Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.
8. All organisms in a test must be from the same source.
9. It is desirable to start tests as soon as possible after collection of sediment from the field.
10. All test chambers and compartments should be identical and should contain the same amount of sediment and overlying water.
11. Negative control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.
12. Culture and test temperatures should be the same. Acclimation of test organisms to the test water is not required.
13. The daily mean test temperature must be within ±1ºC of the desired temperature. The instantaneous temperature must always be within ±3ºC of the desired temperature.
14. Natural physicochemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

Table 16. Test conditions for conducting a 21-d water column toxicity test: *Daphnia magna*.
(Compiled from ASTM 2000h)

Parameter	Conditions
Test type	Static renewal, or flow-through
Temperature	20°C
Light intensity	< 600 lx
Photoperiod	16 h light:8 h dark
Test chamber	500-mL or 2-L screened beaker (flow-through); 100-mL to 1-L covered beaker (static)
Sample volume	30 mL per animal (flow-through); 40 mL per animal (static)
Neonate density	At least 5 per beaker (flow-through); 1+ per beaker (static)
Neonate age	Less than 24 hours
Number of neonates/treatment	10 for single animal per chamber, 20 for multiple animals per chamber
Renewal	3 times weekly (static), one volume daily (flow-through)
Control water	Synthetic or natural water
Water quality	Dissolved Oxygen, pH, conductivity, alkalinity, hardness regularly; temperature daily; ammonia as needed
Cleaning	Three times weekly; clean flow-through screens daily
Feeding	Culture food; daily (static), 2-3 times daily (flow-through)
Aeration	None
Test duration	21 days
Endpoint	Survival, growth and reproduction
Test Acceptability Criteria	70% control survival of first generation, and satisfaction of criteria outlined in Table 17.

Table 17. Test acceptability requirements for a 21-d water column toxicity test with *Daphnia magna*.
(Compiled from ASTM 2000h)

1. Daphnids should be randomly assigned to the test chambers, and there should be at least four chambers or 10 daphnids per treatment.
2. Daphnids should be less than 24 hours old at the start of the test, and should be from a culture exhibiting acceptable reproduction for at least two generations.
3. Appropriate dilution-water and solvent controls should be included in the test.
4. The test should proceed for 21 days.
5. At least 70% of first-generation control daphnids should survive for 21 days.
6. Surviving control daphnids should produce on average at least 60 young in 21 days.
7. Ephippia should be absent from controls.
8. Temperature, dissolved oxygen and pH should be measured as specified.
9. Mean dissolved oxygen in each treatment should be at least 3.0 mg/L, and no dissolved oxygen should be less than 1.5 mg/L.
10. Mean temperature in each treatment should be between 18 and 22ºC, and all temperatures should be between 17 and 23ºC (except for occasional deviations among numerous measurements).

Ceriodaphnia dubia and *Daphnia magna* Toxicity Test Strengths, Limitations and Potential Confounding Factors

The acute and chronic toxicity tests with daphnids *C. dubia* and *D. magna* are among the most widely used in aquatic toxicology. These tests have been subjected to rigorous validation studies as part of their implementation as state and federal regulatory tools. In addition, the ecological relevance of results of these tests as indicators of ecological impacts has been demonstrated in numerous studies. De Vlaming and Norberg-King (1999) reviewed a number of studies where EPA (U.S. EPA 1994c) standardized protocols, including tests with *C. dubia*, were used in conjunction with bioassessment studies, and concluded that these tests are reliable predictors of aquatic ecosystem community responses (see also de Vlaming *et al.* 2001). One additional strength of the tests with *C. dubia* and *D. magna* is that these tests are simple and inexpensive, yet are among the most sensitive of the standardized freshwater toxicity tests available. In addition to the standardized endpoints (survival, fecundity, reproduction) biomarker endpoints have also been incorporated in tests with cladocera, increasing their potential for detecting exposure and sublethal effects (*e.g.*, Day and Scott 1990, Sibley *et al.* 2000).

Another strength of test protocols developed for daphnids is that these tests are amenable to both water column and sediment testing. Tests with both *C. dubia* and *D. magna* have been used extensively in testing of a number of sediment test matrices, including solid-phase, elutriates and porewaters (see review by Burton 1991). In addition Toxicity Identification Evaluation techniques have been shown to be particularly effective with daphnids (see review by de Vlaming *et al.* 2000).

Test results with daphnids may sometimes be influenced by constituents other than anthropogenic chemicals, particularly when testing with sediments and sediment porewater. Examples include unionized ammonia (Ankley *et al.* 1990), manganese (Lasier *et al.* 2000), physical impediments (Sibley *et al.* 1997a) and bicarbonate (Hoke *et al.* 1992). In situations where polycyclic aromatic hydrocarbons are the primary contaminants of concern, the toxicity of these compounds has been shown to increase effects on *D. magna* and *C. dubia* in the presence of UV light (Newsted and Giesy 1987, Ireland *et al.* 1996). Tietge *et al.* (1997) discussed the confounding effects of seven

major ions associated with produced water on tests with *D. magna* and *C. dubia*. Other potential confounding factors are discussed by Winger *et al.* (in review; Table 8).

Freshwater Fish Water Toxicity Tests

Larval Toxicity Tests with *Pimephales promelas* (Fathead Minnow), *Oncorhynchus mykiss* (Rainbow Trout), *Salvelinius fontinalis* (Brook Trout), *Lepomis macrochirus* (Bluegill Sunfish), and *Ictalurus punctatus* (Channel Catfish)

Compiled from ASTM 2000i, U.S. EPA 1993a, U.S. EPA 1994c

Various freshwater fish species are commonly used in water column toxicity testing. *Pimephales promelas* is an omnivorous freshwater fish distributed widely across North America in a variety of habitats, from brooks to small lakes. It is tolerant of high temperature and turbidity, and low oxygen concentrations. Adults of this species reach an average length of 5 cm, and live to about three years of age. *Oncorhynchus mykiss* is an anadromous species native to Pacific coastal streams that has been widely introduced to other parts of North America. Freshwater adults reach an average length of 45 cm and an age of 3 to 4 years. *Salvelinius fontinalis* is native or introduced in much of North America, and can be found in clear, cold brooks, streams, rivers, and small lakes. Adults generally do not exceed 54 cm in length, and 5 years of age. Other freshwater species commonly used in warm-water toxicity tests are *Lepomis macrochirus* and *Ictalurus punctatus*.

Embryos, larvae and fry can be cultured in-house, or obtained from commercial suppliers or trout hatcheries. Larval test organisms are fed prior to and during the test with trout chow (trout species) or *Artemia* nauplii (other species). Test solutions can consist of freshwater samples, freshwater controls, and reference toxicant solutions.

Short-term tests of 1, 2, or 4 days in duration can be performed on all five species. Ten organisms are introduced into each test chamber, and tests are run at 12ºC (trout) or 20 or 25ºC (others) under static or flow-through conditions. Renewals should be performed every 48 hours for static tests. Trout species do not require feeding during these tests. The test endpoint, after 24, 48 or 96 hours of exposure, is survival.

Seven-day tests with *P. promelas* are performed with larvae, to a growth and survival endpoint, or with embryos, to a survival and development endpoint. Ten

embryos or larvae are introduced into each test chamber, and tests are run at 25ºC under static, daily renewal conditions.

Chronic early life-stage tests can be performed with embryos of all five fish species. These tests are performed under flow-through conditions, although the test might be adaptable to static renewal conditions. Fertilized embryos are introduced into individual test cups within larger test chambers. After embryos hatch, larvae are transferred out of the cups and into the test chamber, and are fed regularly with brine shrimp nauplii, trout food, and catfish food, as appropriate. Care must be taken to ensure that the test chambers remain free of excessive detritus and dead organisms. Each fish species has slightly different handling, density, nutrition, and water quality requirements that must be adhered to for test success. Test duration ranges from 32 days post test initiation to 30 days post hatching, after which time the test endpoints, survival and growth, are determined. Survival is determined for embryos, larvae, or overall; growth is determined by larval weight (and sometimes length, for trout species).

Table 18. Test conditions for conducting 24-, 48-, or 96-h water column toxicity tests with freshwater fishes: *Pimephales promelas, Oncorhynchus mykiss, Salvelinius fontinalis, Lepomis macrochirus* and *Ictalurus punctatus*.
(Compiled from U.S. EPA 1993a)

Parameter	Conditions
Test type	Water column static non-renewal, renewal or flow-through
Temperature	*P. promelas, L. macrochirus, I. punctatus*: 20 or 25°C Trout: 12ºC
Light quality	Ambient laboratory illumination
Light intensity	10 to 20 μE/m^2/s
Photoperiod	16 h light:8 h dark
Test chamber	*P. promelas, L. macrochirus, I. punctatus*: 250-mL beaker Trout: 5 L
Test solution volume	*P. promelas, L. macrochirus, I. punctatus*: 200 mL Trout: 4 L
Organism density	10 per beaker
Organism age	*P. promelas, L. macrochirus, I. Punctatus*: 1 to 14 days; 24-h range Rainbow Trout: 15 to 30 days Brook Trout: 30-60 days
Number of replicates/treatment	4
Renewal	48 hours
Dilution/control water	Synthetic water, diluted mineral water, or uncontaminated natural water
Water quality	Temperature continuously; dissolved oxygen, pH daily; conductivity, hardness, alkalinity at start, renewal, and end of test; ammonia and chlorine as needed
Cleaning	None
Feeding	*P. promelas, L. macrochirus, I. punctatus*: *Artemia* nauplii prior to test and renewal
Aeration	None
Test duration	24, 48, or 96 hours
Endpoint	Survival
Test Acceptability Criteria	90% control survival

Table 19. Test conditions for conducting 7-d water column toxicity tests with the freshwater fish *Pimephales promelas*.
(Compiled from U.S. EPA 1994c)

Parameter	Conditions
Test type	Static renewal
Temperature	25°C
Light quality	Ambient laboratory illumination
Light intensity	10 to 20 μE/m^2/s
Photoperiod	16 h light:8 h dark
Test chamber	Survival/growth: 500 mL minimum Survival/teratogenicity: 150 mL minimum
Test solution volume	Survival/growth: 250 mL minimum Survival/teratogenicity: 70 mL minimum
Organism density	15 per beaker (10 minimum)
Organism age	Survival/growth: larvae <24 h or up to 48 h; 24-h range Survival/teratogenicity: embryos <36 h or up to 48 h
Number of replicates/treatment	4 (minimum 3)
Renewal	Daily
Dilution/control water	Synthetic water, diluted mineral water, or uncontaminated natural water
Water quality	Temperature continuously; dissolved oxygen, pH, conductivity, hardness, alkalinity daily; hardness for embryos must be at least 25 mg/L $CaCO_3$
Cleaning	Survival/growth: siphon daily before renewal
Feeding	Survival/growth: *Artemia* nauplii daily, except last 12 hours
Aeration	None
Test duration	7 days
Endpoint	Survival and growth (weight); Survival of normally developed larvae
Test Acceptability Criteria	80% control survival; for growth, average dry weight of control organisms at least 0.25 mg

Table 20. Test conditions for conducting early life stage water column toxicity tests with freshwater fishes: *Pimephales promelas, Oncorhynchus mykiss, Salvelinius fontinalis, Lepomis macrochirus* and *Ictalurus punctatus*. (Compiled from ASTM 2000i)

Parameter	Conditions
Test type	Flow-through (and others)
Temperature	*P. promelas*, *L. macrochirus*, *I. punctatus*: 28ºC Trout: 10ºC
Light quality	Trout: dim light or darkness
Photoperiod	*P. promelas*, *L. macrochirus*, *I. punctatus*: 16 h light:8 h dark
Test chamber	Glass incubation cups (screened) inside stainless steel or glass external chambers (for flow-through)
Test solution volume	Variable
Embryo density	(Depends on species and chamber size) Trout: 1 to 3 per cm^2 (60 per treatment) *P. promelas*: 15 to 20 per large chamber (60 per treatment) *I. punctatus*: 20 embryos in each of two incubation cups per test chamber *L. macrochirus*: 60 per treatment
Embryo age (time after fertilization)	*P. promelas*: 2 to 24 h, <48 h *L. macrochirus*: 2 to 24 h, <48 h *I. punctatus*: within 24 h, < 48 h Trout: within 96 h
Number of replicates/treatment	Variable, depends on test; at least 2 chambers, 40 embryos per treatment
Renewal	At least 5 volumes in 24 hours
Dilution/control water	Synthetic water or uncontaminated natural water
Water quality	Temperature daily;dissolved oxygen, pH, conductivity, hardness, alkalinity at start, weekly, and end of test; ammonia, particulates, chemical oxygen demand (COD), total dissolved gases (TDG) desirable, weekly
Cleaning	Clean or change screens when clogged: clean settled material regularly
Feeding (larvae)	Trout: brine shrimp nauplii and/or trout food, daily *P. promelas*, *L. macrochirus*: brine shrimp nauplii, daily *I. punctatus*: brine shrimp nauplii and catfish food, daily

Table 20 (continued). Test conditions for conducting early life stage water column toxicity tests with freshwater fishes: *Pimephales promelas, Oncorhynchus mykiss, Salvelinius fontinalis, Lepomis macrochirus* and *Ictalurus punctatus*.
(Compiled from ASTM 2000i)

Parameter	Conditions
Aeration	*I. punctatus*: embryos suspended
Test duration	Trout: 30 days post hatching *P. promelas*, *L. macrochirus*, *I. punctatus*: 32 days
Endpoint	Survival (embryos, fry, overall) and growth (weight of survivors, can also measure length for trout)
Test Acceptability Criteria	Control survival (see also Table 21): Trout: 70% survival post-thinning *P. promelas*: 70% survival from 48 h to 32 days *I. punctatus*: 65% survival *L. macrochirus*: 75% mean survival from 48 h to 32 d, and survival in all control chambers at least 65%

Table 21. Test acceptability requirements for a chronic early life-stage toxicity test with freshwater fishes: *Pimephales promelas, Oncorhynchus mykiss, Salvelinius fontinalis, Lepomis macrochirus* and *Ictalurus punctatus*.
(Compiled from ASTM 2000i)

---

1. All chambers and compartments should be identical.
2. Treatments should be randomly assigned to test chamber locations.
3. Required controls should be included.
4. The test should be started with organisms of the appropriate age.
5. The test organisms should be randomly assigned to test chambers and compartments.
6. The test should not be terminated early.
7. Appropriate data on survival and growth must be obtained.
8. Control organisms should survive and grow as specified for each species.
9. Temperature and dissolved oxygen should be measured as specified.
10. The time-weighted average-measured dissolved oxygen concentration should be between 60% and 100% of saturation at the end of the test in all test chambers.
11. The difference between the time-weighted average measured temperatures for two test chambers should be no more than 1ºC.
12. Individual measured temperatures in all chambers should be no more than 3ºC different from the mean of the time-weighted average measured temperatures for the individual test chambers. An exception to this would be if numerous temperature measurements were made, and only one deviation was noted.
13. At any one time, the temperature difference between two test chambers should not be more than 2ºC.

Freshwater Fish Toxicity Test Strengths, Limitations and Potential Confounding Factors

Because of the EPA recommendation that a minimum of three tests representative of invertebrates, algae, and fish, be used, toxicity test protocols for fathead minnows (*P. promelas*) are also among those most commonly used for monitoring and assessment studies. Because fish are key food web components, and are consumed by higher trophic level organisms including humans, tests with fish provide a relevant tool for ERA applications. Tests with *P. promelas* and the other species listed above have proven to be sensitive to a wide variety of chemicals, and are particularly sensitive to certain classes of chemicals such as organochlorine and pyrethroid pesticides, ammonia, and certain metals. The fish species listed above are amenable to laboratory culture, and a number of life stages and endpoints may be used, including embryo-larval development, hatching success, and larval growth and survival. In addition, fish are well suited for studies incorporating biomarker endpoints (Huggett *et al.* 1992). For example, the Japanese medaka (*Orizias latipes*) has been demonstrated to be a useful model for investigating a wide variety of biomarker endpoints (Helmstetter *et al.* 1996). A standardized protocol using *P. promelas* has been developed as a standard for determining the endocrine disruption characteristics of compounds (*e.g.*, Korte *et al.* 1998, Hemming *et al.* 2000). Protocols with the other fish listed above have been demonstrated to be useful for determining toxicity of a wide variety of chemical groups.

Tests with fish are influenced by similar confounding factors as those described for other freshwater species (Table 11). Results of tests with *P. promelas* have been shown to be affected by infestations from fungi and other pathogens (de Vlaming *et al.* 2000). Antibiotics have been used to minimize the influence of pathogens. Tietge *et al.* (1997) discussed the confounding effects of seven major ions associated with produced water on tests with *P. promelas.* In situations where polycyclic aromatic hydrocarbons are the primary contaminants of concern, the toxicity of these compounds has been shown to increase effects on *P. promelas* in the presence of UV light (Oris and Giesy 1987).

Freshwater Amphibian Water Toxicity Tests

African Clawed Frog (*Xenopus laevis*) 96-h Embryo Survival, Growth and Development Toxicity Test (FETAX)

Compiled from ASTM 2000j, Fort pers. comm.

The Frog Embryo Teratogenesis Assay-*Xenopus* (FETAX) is a rapid, sensitive, adaptable developmental toxicity test yielding results that may be extrapolated to mammals and other species. Adults of the test species, *Xenopus laevis*, are obtained from animal suppliers and cultured under controlled conditions until mating is induced by injection of a male and female pair with human chorionic gonadotropin. Eggs are typically fertilized and laid within 12 hours of induction, and embryos immediately checked for quality and life stage.

The 96-h FETAX test is initiated by introducing 25 embryos (stage 8 to 11) into each of two dishes containing 10 mL (or more) of test solution; four dishes are used for controls. Three separate tests are conducted using embryos from three unique mating pairs. Test solutions can consist of freshwater samples, freshwater controls, and reference toxicant solutions, and may be static, flow-through, or renewed daily. The test is run at 24ºC for 96 to 99 hours, depending on the time required for 90% of control embryos to reach the developmental endpoint. Toxicity endpoints include mortality, malformation and growth inhibition.

This test can also be modified for solid-phase samples and their extracts, and for determination of mammalian toxicity through metabolic activation. Solid-phase tests are conducted as sediment-water interface tests, where embryos are submerged in screened tubes to separate them from the sediment layer.

Although the FETAX test was designed specifically for testing with *X. laevis*, the use of endemic, wild-caught frog species might be warranted. The species *Rana catesbiena*, *R. pipiens*, *Bufo americanus* and *B. fowleri* are recommended based on various factors, including egg production, geographical range, and hatching period.

Table 22. Test conditions for conducting a 96-h water column toxicity test for embryonic survival, growth and development: *Xenopus laevis*.
(Compiled from ASTM 2000j, Fort pers. comm.)

Parameter	Conditions
Test type	Water column, static or flow-through, renewal
Temperature	24°C
Light quality	Not applicable
Test chamber	Covered 60-mL glass petri dish
Test solution volume	10 mL (or more to avoid NH_3 buildup)
Embryo density	25/test chamber
Embryo stage	Stage 8 to 11
Number of replicates/treatment	2 per each of three tests (4 for controls)
Dilution/control water	FETAX solution
Water quality	pH 6.5 to 9 optimum; recommended dissolved oxygen at least 40% saturation
Feeding	None
Cleaning	Remove dead embryos daily
Aeration	None
Test duration	96 hours
Endpoint	Survival, growth, development
Test Acceptability Criteria	Negative control survival and normal development: ≥ 90%, and satisfaction of other criteria outlined in Table 23.

Table 23. **Test acceptability requirements for a 96-h water column toxicity test with *Xenopus laevis*.**
(Compiled from ASTM 2000j)

---

1. Embryos from only one mating pair can be used for one test.
2. Metal parts should not be used in breeding aquarium.
3. 90% of FETAX solution controls should reach stage 46 by 99 hours. Low temperature can cause slow development.
4. Dilution water other than FETAX solution should perform similarly to FETAX solution with respect to embryonic growth rate.
5. Any deionized or distilled water used should conform to ASTM Type I standard.
6. Required controls must be included in the test.
7. Embryo staging must be performed according to Nieuwkoop and Faber.
8. Test must be initiated with stage 8 to 11 embryos that are randomly assigned to test chambers.
9. Petri dishes must be physically identical, and randomly distributed (in non-forced-air incubators), within a test.
10. Mortality, growth and development must be properly documented.
11. The pH must remain between 6.5 and 9.0.
12. Dead embryos must be removed every 24 hours.
13. Short temperature deviations greater than 2ºC might be inconsequential, but consistent deviations are not acceptable.
14. Reference toxicant tests must produce means within 2 SD of control chart means.

FETAX Toxicity Test Strengths, Limitations and Potential Confounding Factors

The embryo development test using *Xenopus laevis* represents an important addition to the suite of freshwater protocols available for ERA applications because of the importance of amphibians in freshwater ecosystems, and recent concerns about world-wide amphibian population declines. Although this protocol has been used primarily for assessing toxicity of water samples to date, it has also been used to test sediment samples (Fort *et al.* 1999) and soil extracts. Additional strengths of the FETAX protocol include amenability to including additional biomarker endpoints, adaptation of the protocol for use with multiple species, and thorough interlaboratory comparisons of the protocol. There is insufficient data to determine the relative sensitivity of this test compared to the other tests listed above; most of the reference toxicant testing with this species has emphasized teratogenic compounds (D. Fort, personal communication). There have been few comparisons between FETAX and other standardized protocols using ambient samples.

Results of the FETAX protocol have been shown to be affected by low ionic concentrations in ambient waters (Tietge *et al.* 2000), and possibly by elevated ammonia concentrations. This may be a significant problem in application of this assay for assessing toxicity of ambient samples. The influence on this assay of the other potential confounding factors shown in Table 11 is poorly understood.

Relative Contaminant Sensitivity of Freshwater Water-column Toxicity Tests

The relative sensitivities of the toxicity test protocols used for ambient monitoring can be compared by assessing responses in water-only exposures (*e.g.*, reference toxicant tests), and by comparing side-by-side tests of effluents or ambient samples. Results of water-only exposures were compiled from those reported in the literature (*e.g.*, ECOTOX database and other sources) and in some cases, from unreported experiments conducted at testing laboratories at UC Davis. Results of these experiments are generally reported as 96-h LC50 or EC50 values except where noted. Most of this data can be obtained from the US EPA ECOTOX database (www.epa.gov/ecotox). It is important to note that for both the water column protocols and the sediment protocols discussed below, species responses to contaminants are influenced by a variety of factors including hardness,

alkalinity, conductivity, pH, and temperature. Because of this, the dose-response data summarizing acute toxicity of representative chemicals to freshwater water column and sediment test species often gives ranges of LC50 values for given chemicals. Relative protocol sensitivities to reference chemicals are therefore discussed in general terms. Variation in physico-chemical factors may also affect the relative sensitivities of species and protocols in comparisons using ambient samples.

Water column toxicity test sensitivity to trace metals varies with protocol. Toxicity tests with daphnids (*C. dubia, D. magna*) and the amphipod *H. azteca* were among the most sensitive to copper, while tests with fish were sometimes comparable to these tests under low pH and low hardness conditions. Tests with *B. calyciflorus, L. varigatus, R. subcapitata,* and *C. tentans* were somewhat less sensitive to this metal. Tests with the alga *R. subcapitata,* daphnids and *H. azteca* were sometimes the most sensitive to zinc, while tests with fish, *C. tentans, B. calyciflorus,* and *L. varigatus* were less sensitive. Tests with daphnids, *H. azteca,* and under certain physico-chemical conditions, some fish species, were more sensitive to cadmium. *Raphidocelis* demonstrated moderate sensitivity to this metal, while tests with *L. varigatus, C. tentans, Xenopus laevis,* and *C. tentans* were less sensitive to cadmium.

Sensitivity to organochlorine compounds also varied with species and protocol. In many cases fish species were the most sensitive to pesticides/biocides such as DDT chlordane and pentachlorophenol. For compounds with relatively higher log K_{ow}, sensitivity of species used in sediment exposures may be more relevant. For example, *H. azteca,* and to a lesser extent, *C. tentans and D. magna* were also relatively sensitive to DDT and chlordane, while *L. varigatus* was less sensitive to these compounds. Tests with the alga *R. subcapitata* were comparable to those with some of the fish species in sensitivity to pentachlorophenol, while tests with *Xenopus laevis and B. calyciflorus* were less sensitive to this compound. Tests with *C. tentans, X. laevis,* and a number of fish species were among the most sensitive to dieldrin; *H. azteca* demonstrated moderate sensitivity to this pesticide. A limited amount of relative sensitivity data were available for PCBs. *D. magna* was among the most sensitive invertebrate species to Arochlor 1254, and was comparable to a number of fish species (posted at www.epa.gov/ecotox).

Tests with arthropods and chironomids are the most sensitive to organophosphate and carbamate pesticides. Tests with *C. dubia, C. tentans, H. azteca* and *D. magna* are the most sensitive to chlorpyrifos and diazinon, while tests with fish and algae are considerably less sensitive to these pesticides. Tests with rotifers (*B. calyciflorus*) are the least sensitive to these compounds. Tests with *C. dubia* and *C. tentans* are the most sensitive to the carbamate pesticide carbofuran, while tests with *D. magna* and fish species are less sensitive to this pesticide (posted at www.epa.gov/ecotox). Tests with arthropods and fish are among the most sensitive to the synthetic pyrethroid pesticides. Tests amphipods, and with the daphnids *C. dubia* and *D. magna* are particularly sensitive to the pesticides pyrethrin, fenvalerate and cypermethrin, as are those with the fish species, *L. macrochirus*, *O. mykiss*, and *P. promelas* (Giddings *et al.* 2001). The frog embryo teratogenesis assay with *X. laevis* (FETAX) was less sensitive to permethrin.

Comparative data on freshwater protocol sensitivity to polycyclic aromatic hydrocarbons are limited. Tests with benthic species such as *H. azteca* and *C. tentans* are relatively sensitive to fluoranthene in water-only exposures (posted at www.epa.gov/ecotox). *C. dubia* was more sensitive to this compound than *D. magna*. Sensitivity of protocols using fish species varied depending on the length of exposure. The FETAX protocol demonstrated comparable sensitivity to this compound. Beatty *et al.* (2000) used acute toxicity tests with *C. dubia* (48-h) and the rotifer *B. calyciflorus* (24-h) to assess effects of runoff and receiving water impacted by a petroleum storage and manufacturing site. While *C. dubia* detected significant toxicity in the effluents from this site, the rotifers were unaffected by any of the samples.

Comparisons of the relative sensitivity of standardized test protocols using ambient samples indicate variable sensitivity depending on the contaminants present. Ambient monitoring programs in California were reviewed by de Vlaming *et al.* (2000). In most cases these were comparisons of the daphnid *C. dubia*, the fish *P. promelas*, and the alga, *R. subcapitata*, or in particular instances, the euryhaline mysid *Neomysis mercedis*. In most of these studies, most incidences of toxicity in surface waters were detected by *C. dubia*; TIEs suggested toxicity to this species was most often due to organophosphate pesticides. These authors suggested that where surface waters were toxic to *P. promelas*, this was due either to elevated unionized ammonia or pathogens.

Anderson *et al.* (in review) also found that toxicity in the Calleguas Creek watershed of southern California was most often detected using *C. dubia* compared to *P. promelas* or *R. subcapitata.* When toxicity was detected using *P. promelas*, this was due to elevated concentrations of unionized ammonia, and could be eliminated by reducing the sample pH. In some surface waters of California's Central Valley, toxicity was also detected by the alga, *R. subcapitata.* Analytical chemistry and TIE evidence suggested this was due to elevated metal (Cu and Zn) concentrations associated with mine drainages.

As part of an Ecological Risk Assessment at a superfund site in Tennessee, Suter *et al.* (1999) used the larval fathead minnow growth and survival test, the 7-day survival and fecundity test with *C. dubia*, and embryo-larval development tests with Japanese medaka, redbreast sunfish, and largemouth bass to assess water column toxicity. This site was polluted with a mixture of metals and PCBs. Only the medaka and redbreast sunfish embryo tests showed consistent evidence of toxicity in this study. Toxicity was correlated with elevated metal concentrations (Ni).

Freshwater Whole Sediment Toxicity Tests

Freshwater Amphipod Sediment Toxicity Tests: *Hyalella azteca* 10-d Growth and Survival Test

Compiled from U.S. EPA 2000

*Hyalella azt*eca is a freshwater amphipod species that inhabits lakes, ponds, and streams throughout North and South America. These amphipods are epibenthic detritivores that burrow into the surficial sediment surface (upper 2 cm). Studies with this species have demonstrated they are tolerant of a wide range of sediment physico-chemical characteristics (*e.g.*, variations in grain size, TOC, and conductivity), but are sensitive to chemical contaminants. *Hyalella* has been cultured in reconstituted waters at salinities up to 15 ‰. Several different toxicity tests have been developed with this species; all require daily feeding and renewal of sediment overlying water.

The 10-d toxicity test with *H. azteca* measures survival and growth. This protocol requires 8 replicates each with ten 7-to 14-day-old amphipods (routine testing - modifications to the protocol may be necessary depending on specific research

questions). The amphipods are exposed to 100 mL of sediment in 300 mL beakers, each containing 175 mL of overlying water. The test temperature is 23°C. The overlying water is renewed twice daily, and 1.0 mL of food (Yeast, Cerophyl®, and Trout Chow – YCT) is added to each test container. Research has demonstrated that excessive control mortality may occur if animals are not fed during this test. The containers are not aerated unless dissolved oxygen drops below 2.5 mg/L. If aeration is necessary in any one container, then all test containers are aerated. After 10 days, the amphipods are removed from the sediment by sieving, and the number of surviving animals recorded. Growth is measured using one of two methods. Growth is measured as dry weight per individual amphipod, or as length per animal. Studies have shown that growth is sometimes more sensitive than survival as an indicator of toxicity with *Hyalella*.

Table 24. Test conditions for conducting a 10-d sediment toxicity test with *Hyalella azteca*.
(Compiled from U.S. EPA 2000)

Parameter	Conditions
Test type	Whole-sediment toxicity test with renewal of overlying water
Temperature	23 ± 1ºC
Light quality	Wide-spectrum fluorescent lights
Illuminance	About 100 to 1000 lux
Photoperiod	16 h light:8 h dark
Test chamber	300-mL high-form lipless beaker
Sediment volume	100 mL
Overlying water volume	175 mL
Renewal of overlying water	2 volume additions/d; continuous or intermittent (*e.g.*, 1 volume addition every 12h)
Age of organisms	7- to 14-d old at the start of test (1- to 2-d range in age)
Number of organisms/chamber	10
Number of replicate chambers/treatment	Depends on the objective of the test. Eight replicates are recommended for routine testing
Feeding	YTC food, fed 1.0 mL daily (1800 mg/L stock) to each test chamber.
Aeration	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L.
Overlying water	Culture water, well water, surface water, site water, or reconstituted water
Test chamber cleaning	If screens become clogged during a test, gently brush the *outside* of the screen.
Overlying water quality	Hardness, alkalinity, conductivity, pH, and ammonia at the beginning and end of a test. Temperature and dissolved oxygen daily.
Test duration	10 d
Endpoint	Survival and growth
Test Acceptability	Minimum mean control survival of 80% and measurable growth of test organisms in the control sediment. Additional performance-based criteria specifications are outlined in Table 25

Table 25. Test acceptability requirements for a 10-d Sediment Toxicity Test with *Hyalella azteca*.
(Compiled from U.S. EPA 2000)

A. It is recommended that the following performance criteria be met when conducting a 10-d test with *Hyalella azteca*:

1. Age of *H. azteca* at the start of the test must be between 7- to 14-d old. The 10-d test should start with a narrow range in size or age of *H. azteca* (*i.e.*, 1- to 2-d range in age) to reduce potential variability in growth at the end of the 10-d test.

2. Average survival of *H. azteca* in the control sediment must be greater than or equal to 80% at the end of the test. Growth of the test organisms should be measurable in the control sediment at the end of the 10-d test (*i.e.*, relative to organisms at the start of the test).

3. Hardness, alkalinity, and ammonia in overlying water typically should not vary by more than 50% during the test, and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.

B. Performance-based criteria for culturing *H. azteca* include the following:

1. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicant tests to assess the sensitivity of culture organisms. Data from these reference- toxicant tests could be used to assess genetic strain or life-stage sensitivity of test organisms to selected chemicals.

2. Laboratories should track parental survival in the cultures and record this information using control charts if known-age cultures are maintained. Records should also be kept on the frequency of restarting cultures and the age of the brood organisms.

3. Laboratories should record the following water-quality characteristics of the cultures at least quarterly: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in cultures should be measured weekly. Temperature of cultures should be recorded daily. If static cultures are used, it may be desirable to measure water quality more frequently.

4. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.

5. Physiological measurements such as lipid content might provide useful information regarding the health of cultures.

C. Additional requirements:

1. All organisms in a test must be from the same source.

2. Sediment collected from the field should be stored in the dark at 4ºC and should be held for as little time as possible, though actual recommended storage times vary.

3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.

4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.

5. Test organisms must be cultured and tested at 23°C (±1°C).

6. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.

7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

Freshwater Sediemnt Toxicity Test: *Hyalella azteca* 28-d to 42-d Survival, Growth and Reproduction Toxicity Test

Compiled from U.S. EPA 2000

In addition to the shorter-term growth and survival test with *Hyalella*, a longer-term test that incorporates survival, growth and reproductive endpoints has also been developed. The test is conducted as described above for the 10-d procedure. The two primary differences in experimental design are that the longer-term tests are initiated with 7- to 8-d-old amphipods, and the test starts with 12 replicate containers rather than 8. On day 28, four of the replicate containers are sacrificed and survival and growth are measured as described above. Typically, *Hyalella* are in amplexus (coupling of male and females) from days 21 to day 28 with release of the first brood between day 28 and day 42. The longer-term protocol is designed to incorporate reproduction by transferring the amphipods into clean water (without sediment) so that released young may be counted. Surviving animals from the remaining 8 test containers are separated from the sediment via sieving, and placed in 300 mL beakers with 175 mL of water. A 3 x 3 cm piece of nylon coiled web material is placed into each beaker to provide substrate. Water in these beakers is renewed twice daily and each beaker is fed daily as described above. Reproduction is recorded on day 35 by removing the adults and counting the number of young amphipods in each beaker. The young are removed and the adults are returned to each beaker, then both young and adults are again removed and recounted on day 42. Adult amphipods are preserved, and the number of adult females is determined by counting the number of adult males (those with enlarged second gnathopods), and assuming all other adults are females. The number of adult females is used to determine number of young/female/beaker from Day 28 to Day 42. Growth can also be measured on these adult amphipods.

Table 26. Test conditions for conducting a 42-d sediment toxicity test with *Hyalella azteca*.
(Compiled from U.S. EPA 2000)

Parameter	Conditions
Test type	Whole-sediment toxicity test with renewal of overlying water
Temperature	23 ± 1ºC
Light quality	Wide-spectrum fluorescent lights
Illuminance	About 100 to 1000 lux
Photoperiod	16 light:8 dark
Test chamber	300-mL high-form lipless beaker
Sediment volume	100 mL
Overlying water volume	175 mL in the sediment exposure from Day 0 to Day 28 (175 to 275 mL in the water-only exposure from Day 28 to Day 42)
Renewal of overlying water	2 volume additions/d; continuous or intermittent (*e.g.*, 1 volume addition every 12h)
Age of organisms	7- to 8-d old at the start of test
Number of organisms/chamber	10
Number of replicate chambers/treatment	12 (4 for 28-d survival and growth; and 8 for 35- and 42-d survival, growth and reproduction). Reproduction is more variable than growth or survival; hence more replicates might be needed to establish statistical differences among the treatments.
Feeding	YTC food, fed 1.0 mL daily (1800 mg/L stock) to each test chamber.
Aeration	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L.
Overlying water	Culture water, well water, surface water, or site water. Use of reconstituted water is not recommended.
Test chamber cleaning	If screens become clogged during a test, gently brush the *outside* of the screen.
Overlying water quality	Hardness, alkalinity, conductivity, and ammonia at the beginning and end of a sediment exposure (Day 0 and 28). Temperature daily. Conductivity weekly. Dissolved oxygen (DO) and pH three times/ week. Concentrations of DO should be measured more often if DO drops more than 1 mg/L since the previous measurement.

Table 26. Test conditions for conducting a 42-d sediment toxicity test with *Hyalella azteca*.
(Compiled from U.S. EPA 2000)

Parameter	Conditions
Test duration	42 d
Endpoint	28-d survival and growth; 35-d survival and reproduction; and 42-d survival, growth, reproduction, and number of adult males and females on Day 42.
Test Acceptability	Minimum mean control survival of 80% on Day 28. Additional performance-based criteria specifications are outlined in Table 27 based on results of round-robin testing.

Table 27. Test acceptability requirements for a 42-d sediment toxicity test with *Hyalella azteca*.
(Compiled from U.S. EPA 2000)

A. It is recommended that the following performance criteria be met when conducting a 10-d test with *Hyalella azteca*:

1. Age of *H. azteca* at the start of the test should be between 7- to 8-d old. Starting a test with younger or older organisms may compromise the reproductive endpoint.

2. Average survival of *H. azteca* in the control sediment on Day 28 should be greater than or equal to 80%.

3. Laboratories participating in round-robin testing reported after 28-d sediment exposure in a control sediment, survival >80% for >88% of the laboratories; length >3.2 mm/individual for >71% of the laboratories; and dry weight >0.15 mg/individual for >66% of the laboratories. Reproduction from Day 28 to Day 42 was >2 young/female for >71% of the laboratories participating in the round-robin testing. Reproduction was more variable within and among laboratories; hence, more replicates might be needed to establish statistical differences among treatments with this endpoint.

4. Hardness, alkalinity, and ammonia in overlying water typically should not vary by more than 50% during the sediment exposure, and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.

B. Performance-based criteria for culturing *H. azteca* include the following:

1. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms. Data from these reference-toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to selected chemicals.

2. Laboratories should track parental survival in the cultures and record this information using control charts if known-age cultures are maintained. Records should also be kept on the frequency of restarting cultures and the age of the brood organisms.

3. Laboratories should record the following water-quality characteristics of the cultures at least quarterly: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in cultures should be measured weekly. Temperature of cultures should be recorded daily. If static cultures are used, it may be desirable to measure water quality more frequently.

Table 27 (continued). Test acceptability requirements for a 42-d sediment toxicity test with *Hyalella azteca*.
(Compiled from U.S. EPA 2000)

---

4. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.

5. Physiological measurements such as lipid content might provide useful information regarding the health of cultures.

C. Additional requirements:

1. All organisms in a test must be from the same source.

2. Sediment collected from the field should be stored in the dark at 4ºC and should be held for as little time as possible, though actual recommended storage times vary.

3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.

4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.

5. Test organisms must be cultured and tested at 23°C (±1°C).

6. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.

7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

Hyalella azteca Sediment Toxicity Test Strengths, Limitations and Potential Confounding Factors

Compiled from U.S. EPA 2000

Acute and chronic test procedures developed for the freshwater amphipod *H. azteca* are among the most commonly used protocols in freshwater toxicity assessment studies. A review of the sensitivity of protocols with *H. azteca* using water-only exposure data and results of studies with ambient sediment samples show that protocols with this species are often among the most sensitive available for sediment quality assessments. This protocol has been subjected to rigorous interlaboratory testing, and is amenable to *in situ* exposures (*e.g.*, Burton *et al.* 2000). This species is tolerant of sediment and water-only exposures, so it is also amenable to a variety of TIE manipulations (Ankley *et al.* 1991a). Because this species has a relatively short generation time and is easily cultured, a number of endpoints may be assessed (*e.g.*, survival, growth, and reproduction). These provide more flexibility for application in the sediment risk assessment process, and may be applied to population biology models (*e.g.*, reproduction – DeWitt *et al.* 1997). Canfield *et al.* (1994) demonstrated the ecological relationship between toxicity detected with *H. azteca*, and impacts on macrobenthic populations (*e.g.*, increased Chironomidae genera richness). The ecological relationship between toxicity detected by tests with *H. azteca* in laboratory toxicity tests, and impacts on macrobenthic community structure are summarized in EPA (U.S. EPA 2000). Anderson *et al.* (manuscript in preparation) found a positive relationship between increased survival of *H. azteca* in laboratory exposures, greater macrobenthic species richness in field samples, and greater numbers of sensitive indicator species in a California river impacted by agricultural drain water. One other positive attribute of *Hyalella* is that this genus is found in California freshwater habitats (personal communication, P. Ode, California Department of Fish and Game).

H. azteca is relatively tolerant of a wide range of sediment parameters, including grain size and TOC (U.S. EPA 2000). Results of tests with *H. azteca* have been shown to be influenced by elevated manganese (Lasier *et al.* 2000), and alkalinity (Lasier *et al.* 1997), and may be affected by high unionized ammonia concentrations (Ankley *et al.* 1990) and sediments with high proportions of silt/clay. In addition, tests with *H. azteca*

may be influenced by the presence of indigenous organisms. In situations where polycyclic aromatic hydrocarbons are the primary contaminants of concern, the toxicity of these compounds may increase effects on *H. azteca* in the presence of UV light.

Freshwater Sediment Toxicity Tests with Other Invertebrates

*Chironomus tentans*10-d Survival and Growth Toxicity Test

Compiled from U.S. EPA 2000

The midge, *Chironomus tentans* is a dipteran fly species whose larval and pupal stages are found in eutrophic ponds and lakes in northern latitudes. The life cycle of *C. tentans* can be divided into four stages: (1) an egg stage (~ 3 d), (2) a larval stage, consisting of four instars (~ 18 d), (3) a pupal stage (~3 d) , and (4) and adult stage (emergent; ~3 to 5 d). The larval stages of *C. tentans* occur in the upper few cm of the sediment, are tolerant of a number of sediment physico-chemical properties, and are relatively sensitive to contaminants. The relatively short generation time of this species makes it amenable to both short-term and life-cycle laboratory toxicity testing.

The 10-d test with *C. tentans* is started with 10 second- to third-instar larvae (~ 10 d old). The test conditions, containers, sediment, and overlying water volumes, and renewal rates are the same as those described above for *H. azteca*. The larvae are fed 1.5 mL of a 4 g/L Tetrafin® suspension daily. During the test the larvae burrow into the sediment and construct tubes (cases). They feed on particulate matter drawn into tubes or in the vicinity of either end of the open-ended tubes. After 10 d the surviving animals are sieved from the sediment, and survival is recorded. Growth may be measured as length, or dry weight. If dry weight is used, animals are measured as ash-free dry weight (AFDW). AFDW is used because Sibley *et al.* (1997b) found that sediment grain size distributions influence the amount of sediment *C. tentans* larvae ingest and retain in their gut. As a result, a substantial portion of larval weight may be comprised of sediment rather than tissue when testing finer grained sediments. This may confound interpretation of test results (Note: this has not been found to be a problem with *Hyalella* because sediment apparently does not comprise a large proportion of the overall dry weight with this species).

Table 28. Recommended Test Conditions for conducting a 10-d Sediment Toxicity Test with *Chironomus tentans*.
(Compiled from U.S. EPA 2000)

Parameter	Conditions
Test type	Whole-sediment toxicity test with renewal of overlying water
Temperature	23 ± 1ºC
Light quality	Wide-spectrum fluorescent lights
Illuminance	About 100 to 1000 lux
Photoperiod	16 h light:8 h dark
Test chamber	300-mL high-form lipless beaker
Sediment volume	100 mL
Overlying water volume	175 mL
Renewal of overlying water	2 volume additions/d; continuous or intermittent (*e.g.*, 1 volume addition every 12h)
Age of organisms	Second- to third-instar larvae (about 10-d-old larvae; all organisms must be third instar or younger with at least 50% of the organisms at the third instar).
Number of organisms/chamber	10
Number of replicate chambers/treatment	Depends on the objective of the test. Eight replicates are recommended for routine testing
Feeding	Tetrafin® goldfish food, fed 1.5 mL daily to each test chamber (1.5 mL contains 6.0 mg of dry solids)
Aeration	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L.
Overlying water	Culture water, well water, surface water, site water, or reconstituted water
Test chamber cleaning	If screens become clogged during a test, gently brush the *outside* of the screen.
Overlying water quality	Hardness, alkalinity, conductivity, pH, and ammonia at the beginning and end of a test. Temperature and dissolved oxygen daily.
Test duration	10 d
Endpoint	Survival and growth (ash-free dry weight (AFDW))
Test Acceptability	Minimum mean control survival must be 70%, with minimum weight/surviving control organism of 0.48 mg AFDW. Performance-based criteria specifications are contained in Table 29.

Table 29. Test Acceptability Requirements for a 10-d Sediment Toxicity Test with *Chironomus tentans.*
(Compiled from U.S. EPA 2000)

A. It is recommended that the following performance criteria be met when conducting a 10-d test with *Chironomus tentans*:

1. Tests should be started with second- to third-instar larvae (about 10-d-old larvae).

2. Average survival of *C. tentans* in the control sediment must be greater than or equal to 70% at the end of the test.

3. Average size of *C. tentans* in the control sediment must be at least 0.48 mg AFDW at the end of the test.

4. Hardness, alkalinity, and ammonia in overlying water typically should not vary by more than 50% during the test, and dissolved oxygen in the overlying water should be maintained above 2.5 mg/L.

B. Performance-based criteria for culturing *C. tentans* include the following:

1. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicant tests to assess the sensitivity of culture organisms. Data from these reference-toxicant tests could be used to assess genetic strain or life-stage sensitivity of test organisms to selected chemicals.

2. Laboratories should keep a record of time to first emergence for each culture and record this information using control charts. Records should also be kept on the frequency of restarting cultures.

3. Laboratories should record the following water-quality characteristics of the cultures at least quarterly: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in cultures should be measured weekly. Temperature of cultures should be recorded daily. If static cultures are used, it may be desirable to measure water quality more frequently.

4. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.

5. Physiological measurements such as lipid content might provide useful information regarding the health of cultures.

Table 29 (continued). Test acceptability requirements for a 10-d sediment toxicity test with *Chironomus tentans.*
(Compiled from U.S. EPA 2000)

C. Additional requirements:

1. All organisms in a test must be from the same source.

2. Sediment collected from the field should be stored in the dark at 4ºC and should be held for as little time as possible, though actual recommended storage times vary.

3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.

4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.

5. Test organisms must be cultured and tested at 23°C (±1°C).

6. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.

7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

Chironomus tentans Lifecycle Toxicity Test

Compiled from U.S. EPA 2000

The midge *Chironomus tentans* is amenable to both shorter-term and life cycle toxicity testing because it has a relatively short generation time (~ 25 to 30 d at 23 ºC). The life-cycle test with this species is started with newly hatched larvae (< 24 h old), and continues through emergence and reproduction of the adults, and hatching of the F_1 generation. Survival is determined at 20 d and at the end of the test (~ 50 to 65 d). Growth is determined at 20 d. From day 23 to the end of the test, emergence and reproduction are monitored daily. The number of eggs is determined for each egg case, which is incubated for 6 days to determine hatching success. Each treatment of the life cycle test is ended separately when no additional emergence has been recorded for 7 consecutive days.

For routine testing, the test is started with 16 replicates, each with 12 larvae (test containers and test conditions are the same as those described above for the 10-d test). Four replicates are used for the 20-d survival and growth endpoints, and 8 replicates are used for determination of emergence and reproduction. Because *C. tentans* males typically begin emerging 4 to 7 d before the females, additional (auxiliary) males need to be present during the prime female emergence period. The auxiliary males are provided in the 4 additional replicates. On day 30, emergence traps are placed on the 4 additional replicates to collect additional males for use with the females emerging from the reproductive replicates. Males from a different replicate within the same sediment treatment may be paired with females of replicates where no males have emerged. Pairing occurs when adults are transferred to reproduction/oviposit chambers. Females usually oviposit a single primary egg case within 1 d of fertilization. These egg cases are transferred to a petri dish, where egg density is estimated and hatching success monitored. Hatching success is quantified after 6 d of incubation at 23 ºC. Because reproductive output may be quantified as number of eggs per female, hatching success is considered a discretionary endpoint used in studies that require demographic parameters. In summary, this procedure incorporates a lethal and several possible sublethal endpoints: survival (20-d), growth (20-d), emergence (23-d and on), reproduction as number of eggs per female (23-d and on), and percent hatching success (23-d and on).

Table 30. Recommended test conditions for conducting a long-term sediment toxicity test with *Chironomus tentans*.
(Compiled from U.S. EPA 2000)

Parameter	Conditions
Test type	Whole-sediment toxicity test with renewal of overlying water
Temperature	23 ± 1ºC
Light quality	Wide-spectrum fluorescent lights
Illuminance	About 100 to 1000 lux
Photoperiod	16 h light:8 h dark
Test chamber	300-mL high-form lipless beaker
Sediment volume	100 mL
Overlying water volume	175 mL
Renewal of overlying water	2 volume additions/d; continuous or intermittent (*e.g.*, 1 volume addition every 12h)
Age of organisms	< 24-h-old larvae
Number of organisms/chamber	12
Number of replicate chambers/treatment	16 (12 at day–1 and 4 for auxiliary males on day 10)
Feeding	Tetrafin® goldfish food, fed 1.5 mL daily to each test chamber starting day -1 (1.0 mL contains 4.0 mg of dry solids)
Aeration	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L.
Overlying water	Culture water, well water, surface water, site water, or reconstituted water
Test chamber cleaning	If screens become clogged during a test, gently brush the *outside* of the screen.
Overlying water quality	Hardness, alkalinity, conductivity, and ammonia at the beginning, on Day 20, and at end of a test. Temperature daily (ideally continuously). Dissolved oxygen (DO) and pH three times/week. Conductivity weekly. Concentrations of DO should be measured more often if DO has declined by more than 1 mg/L since previous measurement.
Test duration	About 50 to 65 d: each treatment is ended separately when no additional emergence has been recorded for seven consecutive days. When no emergence is recorded from a treatment, termination of that treatment should be based on the control sediment using this 7-d criterion.

Table 30 (continued). Recommended test conditions for conducting a long-term sediment toxicity test with *Chironomus tentans*.
(Compiled from U.S. EPA 2000)

Parameter	Conditions
Endpoint	20-d survival and weight; female and male emergence, adult mortality, the number of egg cases oviposited, the number of eggs produced, and the number of hatched eggs. Potential sublethal endpoints are listed in Table 32
Test Acceptability	Average size of *C. tentans* in the control sediment at 20 d must be at least 0.6 mg/ surviving organism as ash-free dry weight. Emergence should be greater than or equal to 50%. Experience has shown that pupae survival is typically >83% and adult survival is >96%. Time to death after emergence is <6.5d for males and <5.1 d for females. The mean number of eggs/ egg case should be greater than or equal to 800 and the percent hatch should be ≥80%.

Table 31. Test acceptability requirements for a long-term sediment toxicity test with *Chironomus tentans.*
(Compiled from U.S. EPA 2000)

A. It is recommended that the following performance criteria be met when conducting a long-term test with *Chironomus tentans*:

1. Tests must be started with less than 1-d-old larvae (<24-h). Starting a test with substantially older organisms may compromise the emergence and reproductive endpoint.

2. Average survival of *C. tentans* in the control sediment should be greater than or equal to 70% on Day 20 and greater than 65% at the end of the test.

3. Average size of *C. tentans* in the control sediment at 20 d must be at least 0.6 mg/surviving organism as AFDW. Emergence should be greater than or equal to 50%. Experience has shown that pupae survival is typically >83% and adult survival is >96%. Time to death after emergence is <6.5d for males and <5.1 d for females. The mean number of eggs/ egg case should be greater than or equal to 800 and the percent hatch should be ≥80%.

4. Hardness, alkalinity, and ammonia in overlying water typically should not vary by more than 50% during the test, and dissolved oxygen in the overlying water should be maintained above 2.5 mg/L.

B. Performance-based criteria for culturing *C. tentans* include the following:

1. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicant tests to assess the sensitivity of culture organisms. Data from these reference-toxicant tests could be used to assess genetic strain or life-stage sensitivity of test organisms to selected chemicals.

2. Laboratories should keep a record of time to first emergence for each culture and record this information using control charts. Records should also be kept on the frequency of restarting cultures.

3. Laboratories should record the following water-quality characteristics of the cultures at least quarterly: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in cultures should be measured weekly. Temperature of cultures should be recorded daily. If static cultures are used, it may be desirable to measure water quality more frequently.

4. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.

Table 31 (continued). Test acceptability requirements for a long-term sediment toxicity test with *Chironomus tentans.*
(Compiled from U.S. EPA 2000)

---

5. Physiological measurements such as lipid content might provide useful information regarding the health of cultures.

C. Additional requirements:

1. All organisms in a test must be from the same source.

2. Sediment collected from the field should be stored in the dark at 4ºC and should be held for as little time as possible, though actual recommended storage times vary.

3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.

4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.

5. Test organisms must be cultured and tested at 23°C (±1°C).

6. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.

7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

Table 32. Endpoints for a long-term sediment toxicity test with *Chironomus tentans*. (Compiled from U.S. EPA 2000)

Lethal	**Sublethal**		
Survival	Growth	Emergence	Reproduction
Larvae (20-d)	Larvae	Total/Percent	Sex Ratio
Larvae (End)		Cumulative (Rate)	Time to Oviposition
Pupae		Time to first	Mean Eggs/Female
Adults		Time to Death	Egg
			Cases/Treatment
			Egg Hatchability

Chironomus tentans (*C. riparius*) Sediment Toxicity Test Strengths, Limitations and Potential Confounding Factors

Compiled from U.S. EPA 2000

Like *H. azteca*, acute and chronic protocols developed for the midge *Chironomus tentans* are among those most commonly used in freshwater sediment quality assessments. This protocol has been subjected to rigorous interlaboratory testing, and is amenable to *in situ* exposures (*e.g.*, Sibley *et al.* 1999). Acute and chronic protocols developed with *C. tentans* include a number of endpoints including, survival, growth, emergence and various reproductive endpoints. These increase the flexibility for application of these procedures in site-specific ecological risk assessments. *Chironomus* has also been shown to be useful for assessing biomarker endpoints (*e.g.*, acetylcholinesterase inhibition; Fisher *et al.* 2000). The ecological relationship between toxicity detected by tests with *C. tentans* in laboratory toxicity tests, and impacts on macrobenthic community structure are summarized in EPA (U.S. EPA 2000).

Results of tests with *C. tentans* may be influenced by sediment organic matter (Lacey *et al.* 1999) and unionized ammonia. In addition, sediment in the gut may influence the growth endpoint at termination. To eliminate effects of gut sediment, ash-free dry weight is recommended as the index for growth. Protocols with this species have been shown to be influenced by indigenous organisms (U.S. EPA 2000). *Chironomus tentans* are apparently not as easily cultured as *H. azteca*, so this test species may not be as readily available (P. Winger, personal communication).

Hexagenia limbata 10-d Toxicity Test

Mayfly nymphs (Ephemeroptera) are common in soft mud and fine silt/clay benthic habitats of lakes, rivers, and ponds of the United States and Canada. Nymphs of *Hexagenia limbata* (Ephemeroptera: Ephemeridae) have been used in sediment toxicity assessments since the 1970's, and have proven to be relatively sensitive to contaminants. *Hexagenia* nymphs or eggs are collected from the field for testing because the life cycle of this species is too long to be practical for continuous laboratory culturing. Adults are obtained during emergence periods in the summer months and females are placed on the water surface to induce egg extrusion. The eggs are transferred to the laboratory and

warmed incrementally to promote development. Hatching begins in 6 to 8 days at 20 °C. Soon after hatching the nymphs are place into aquaria for rearing. Nebeker *et al.* (1984) described a 10-d, static toxicity test using 10 young *H. limbata* nymphs (< 10mm long) placed into 1 L beakers containing 200 mL of sediment and 800 mL of overlying water. Mortality was measured after 10 d.

Hexagenia limbata 21-d Toxicity Test

Bedard *et al.* (1992) described a 21-d, static whole sediment toxicity test using 10 early instar mayfly nymphs (< 8 mm long) placed into 1.8 L jars (minimum 3 replicates). Each jar held 325 mL of sediment and 1300 mL of overlying water (4: 1 (v:v) water: sediment ratio). Each replicate was aerated. Food was not provided. On day 21, surviving nymphs were separated from the sediment via sieving, survival and growth was determined.

Hexagenia limbata Sediment Toxicity Test Strengths, Limitations and Potential Confounding Factors

Compiled from U.S. EPA 2000

Relative to EPA's criteria listed above (See U.S. EPA, 2000: Table 1.3), the mayfly genus *Hexagenia* has several desirable characteristics, including sensitivity to contaminants, contact with the sediment, ecological importance, and confirmative studies indicating that laboratory response with this species is correlated with impacts on benthic populations (U.S. EPA 2000). A number of studies have demonstrated the utility of this species for sediment toxicity assessments. One of the primary limitations associated with *H. limbata* and other mayfly species is that they have not been demonstrated to be amenable to laboratory culture and therefore may not be tested year-round. In addition, studies have indicated that *H. limbata* may be sensitive to sediment physico-chemical characteristics such as grain size (discussed above). Protocols with *H. limbata* have not been subjected to sufficient round-robin testing, and there is not an adequate database demonstrating the relative sensitivity of this protocol using water-only exposures and field sediment samples contaminated with complex chemical mixtures.

Relative Contaminant Sensitivity of Freshwater Sediment Toxicity Tests

In many freshwater sediment studies, the same species are used in both porewater and solid-phase exposures, allowing for direct comparisons of their relative sensitivity. For example, Giesy *et al.* (1990) compared porewater exposures using *D. magna* (48-h survival) to whole-sediment and porewater exposures using *C. tentans* (10-d growth and survival) and the mayfly *H. limbata* (7-d survival). In this study, *H. limbata* exposed to porewater was the most sensitive assay. This test was only slightly more sensitive than *D. magna* exposed to porewater, or *C. tentans* growth. Cairns *et al.* (1984) compared the relative sensitivity of tests with four species to copper-spiked sediments. In this study *D. magna* was more sensitive than *Chironomus* or the amphipods *Hyalella* and *Gammarus*. Ankley *et al.* (1991b) evaluated the acute toxicity of whole-sediment, porewater, and sediment elutriate samples using *P promelas*, *C. dubia*, *H. azteca*, and *L. varigatus*. These authors found that the amphipod *H. azteca* was the most sensitive species in both porewater and bulk-phase exposures. The test with *P. promelas* was less sensitive than tests with *C. dubia* or *H. azteca.* The test with the oligochaete *L. varigatus* was the least sensitive of those tested.

Sibley *et al.* (1997a) investigated sediment toxicity associated with a pulp mill discharge using a number of acute and chronic test procedures in both porewater and bulk-phase exposures. Acute (48-h) bulk-phase and porewater exposures were conducted with *H. limbata*, *D. magna*, *C. riparius*, *H. azteca*, and the oligochaete worm *Tubifex tubifex*. Chronic tests were conducted with *H. azteca* (28-d growth and survival), *C. riparius* (10-d growth and survival), *D. magna* (30-d reproduction), and *T. tubifex* (30-d reproduction). *H. limbata* was the most sensitive species used in the acute bulk-sediment tests, although this may have been partly due to low dissolved oxygen or grain size. *H. limbata* was also the most sensitive species in the acute porewater exposures, and *T. tubifex* was the least sensitive species in the acute porewater exposures. The test with *D. magna* was comparable to those with *H. azteca* and *C. riparius* in acute bulk-sediment and porewater exposures, although these authors suggested that *D. magna* mortality in the bulk-phase exposures may have been caused by entrapment in organic debris. No effect on reproduction was observed in the chronic tests in this study. Growth of the amphipods and chironomids was stimulated in this study, apparently by organic enrichment of the sediments.

Bay *et al.* (in review) reviewed the relative performance of freshwater bulk-phase and porewater toxicity test protocols, and found that, in most cases, toxicity was more often detected

in porewater exposures than bulk-phase exposures, even when the same species was tested in both test matrices. These authors also evaluated test performance in terms of the amount of unique toxicity information produced by each test method (porewater vs. bulk-phase) Evidence of unique toxicity information was defined as the percentage of the total number of toxic samples detected by either test, that were classified as toxic by only one test method (either porewater or bulk-phase). In the case of freshwater test methods, these authors found that porewater tests provided most or all of the unique toxicity information regardless of the species tested. However, this pattern was reversed for bulk-phase tests incorporating chronic or sublethal endpoints.

Depending on the toxicants present, use of chronic tests incorporating sublethal endpoints can improve the sensitivity of freshwater sediment tests. EPA (U.S. EPA 2000) has reviewed studies where sublethal endpoints provided greater sensitivity relative to survival in bulk-phase exposures. In highly contaminated sediments tested with the amphipod *Hyalella azteca,* survival or growth endpoints identified a similar percentage of samples as toxic in both 14- and 28-d bulk-phase toxicity tests. In some cases, moderately contaminated sediments inhibit growth but not survival. For example, in 28-d tests with *H. azteca,* Kemble *et al.* (1994) found 13% of samples from the Clark Fork River inhibited survival, while 53% inhibited amphipod growth. In some cases, no extra information is obtained with sublethal endpoints using this species (*e.g.*, Day *et al.* 1995). Kemble *et al.* (2000) compared the relative sensitivity of survival to sublethal endpoints using *H. azteca* and *C. tentans* exposed to cadmium. In 42- to 60-d water-only exposures using *H. azteca,* survival at 28-d was a more sensitive endpoint than growth or reproduction, while experiments with *C. tentans* indicated that sublethal endpoints were more sensitive than survival. In this study, reproduction of midges and amphipods was only reduced at cadmium concentrations that also inhibited growth.

Freshwater Sediment Toxicity Test Strengths, Limitations and Potential Confounding Factors

Because the toxicity test protocols discussed below have been used to test bulk-phase and porewater sediment matrices, the strengths, limitations, and influence of confounding factors on these protocols will be discussed in the context of testing both sediment matrices. Note that although these are the test protocols most commonly used for freshwater sediment testing (U.S. EPA 2000), some of the water column protocols discussed above are also used to test both bulk-phase and porewater samples. For example, because of their ecological significance, proximity to epibenthic habitats, and sensitivity to contaminants, cladocerans such as *C. dubia* and *D. magna* are particularly appropriate for sediment assessments. The strengths, limitations and confounding factors associated with these protocols have been discussed above.

Relative to the criteria used by Ingersoll *et al.* (1995a), for selecting appropriate tests to assess the ecological risk of contaminated sediments, acute and chronic toxicity tests with *H. azteca* and *C. tentans* provide the least uncertainty. EPA (See U.S. EPA 2000; Table 1.3) summarizes the criteria for selection of appropriate species for freshwater sediment quality assessments. Criteria include relative sensitivity, evaluation with round-robin studies, species contact with the sediment, taxonomic identification, ecological importance, geographical distribution, tolerance to sediment physico-chemical characteristics, correlation between laboratory response endpoints and impacts on benthic populations, and peer reviewed methods. The test protocols developed for *H. azteca* and *C. tentans* meet all of these criteria. The test with *Lumbriculus varigatus* met all of the criteria except round-robin testing. Information specific to *H. azteca, C. tentans, L. varigatus* and *H. limbata* are found in those sections.

FRESHWATER SEDIMENT POREWATER TESTS

All of the whole sediment test species described above have also been used in porewater exposures (see review by Burton 1991). Toxicity test methods for porewater exposures generally follow those described above for the various whole-sediment protocols except that exposure times are abbreviated and tests are usually conducted in

small volume containers. Porewater is extracted via centrifugation (Ankley *et al.* 1991a), or some times with a syringe extractor (Winger and Lasier 1991).

Giesy *et al.* (1990) used a number of freshwater species in porewater exposures, including *Daphnia magna* (48-h), *Hexagenia limbata* (7-d), *Chironomus tentans* (10-d), and Microtox (*Photobacterium phosphoreum;* 15-min.). In most cases, porewaters were renewed during these tests to maintain water quality and minimize loss of contaminants. Results of these tests were compared to tests using the same species exposed to whole sediment. Ankley *et al.* (1991b) conducted similar comparisons using fathead minnow larvae (*Pimephales promelas*, 96-h), daphnids, *Ceriodaphnia dubia*, 48-h), amphipods (*Hyalella azteca*, 96-h), and oligochaetes (*Lumbriculus varigatus*). Sibley *et al.* (1997a) compared toxicity of sediments and porewaters using five macroinvertebrates, *Hyalella azteca, Daphnia magna, Chironomus riparius, Hexagenia spp*., and *Tubifex tubifex*. Tests were conducted in small volume centrifuge tubes or beakers (40-60 mL porewater) using 48- and 96-h exposures, depending on species. More recently, Winger *et al.* (2000) used the amphipod *Hyalella azteca* in synoptic porewater and whole-sediment exposures. The relative sensitivities of these protocols to sediments and porewaters will be discussed in succeeding sections.

Freshwater Sediment Porewater Toxicity Test Strengths, Limitations and Potential Confounding Factors

See sections on specific test protocols

FRESHWATER SEDIMENT BIOACCUMULATION TESTS

Sediment bioaccumulation using the oligochaete, *Lumbriculus varigatus*, 28 days

Compiled from U.S. EPA 2000

The oligochaete *Lumbriculus varigatus* is recommended for freshwater sediment bioaccumulation testing because it is a deposit-feeding worm that lives in close contact with sediment but is tolerant of a wide range of sediment physico-chemical characteristics. In addition, this species is less sensitive to contaminants, and therefore may tolerate the longer-term (*e.g.*, 28-d) exposures necessary to assess steady-state

bioaccumulation. This species is amenable to laboratory culture so known age individuals may be produced for year-round toxicity and bioaccumulation testing.

The testing procedures for *L. varigatus* are similar to those described previously for marine sediment bioaccumulation tests. Before starting a 28-d bioaccumulation test with *L. varigatus*, toxicity of the test sediment is assessed using a 4-d exposure (4 replicates; 10 animals per replicate). This screening toxicity test is used to verify that the sediment is not acutely toxic to oligochaetes, and that the worms do not avoid the sediment. The number of *L. varigatus* in the 4-d toxicity screening test should not be significantly reduced in the test sediment relative to the control sediment.

The 28-d sediment bioaccumulation test with *L. varigatus* is conducted with adult oligochaetes at 23 °C with a 16 h light:8 h dark photoperiod. Test chambers can be 4 to 6 L containers with 1 to 2 L of sediment and 1 to 4L of overlying water. A minimum of 5 replicates is recommended for routine testing although the number of replicates depends on the objectives of the test. To minimize depletion of sediment contaminants during the course of the test, the ratio of total organic carbon in sediment to dry weight of organisms should be about 50:1. A minimum of 1g sediment/replicate with up to 5g sediment/replicate should be tested. The oligochaetes are not fed during the test. Each replicate receives a minimum of 2 volume additions/d of overlying water to maintain water quality. At the end of the exposure, the test animals are sieved from the sediment, and transferred to clean water to allow their guts to be purged (6 to 8h). Total lipids are measured on a subsample of the total tissue mass from each replicate sample.

Table 33. Recommended test conditions for conducting a 28-d sediment bioaccumulation test with *Lumbriculus varigatus*.
(Compiled from U.S. EPA 2000)

Parameter	Conditions
Test type	Whole-sediment bioaccumulation test with renewal of overlying water
Temperature	23 ± 1ºC
Light quality	Wide-spectrum fluorescent lights
Illuminance	About 100 to 1000 lux
Photoperiod	16 h light:8 h dark
Test chamber	4- to 6-L aquaria with stainless steel screens or glass standpipes
Sediment volume	1 L or more depending on TOC
Overlying water volume	1 L or more depending on TOC
Renewal of overlying water	2 volume additions/d; continuous or intermittent (*e.g.*, 1 volume addition every 12h)
Age of organisms	Adults
Loading of organisms in chamber	Ratio of total organic carbon in sediment to organism dry weight should be no less than 50:1. Minimum of 1 g/replicate. Preferably 5 g/replicate.
Number of replicate chambers/treatment	Depends on the objective of the test. Five replicates are recommended for routine testing.
Feeding	None
Aeration	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L.
Overlying water	Culture water, well water, surface water, site water, or reconstituted water
Test chamber cleaning	If screens become clogged during a test, gently brush the *outside* of the screen.
Overlying water quality	Hardness, alkalinity, conductivity, pH, and ammonia at the beginning and end of a test. Temperature and dissolved oxygen daily.
Test duration	28 d
Endpoint	Bioaccumulation
Test Acceptability	Performance-based criteria specifications are outlined in Table 34.

Table 34. Test acceptability requirements for a 28-d sediment bioaccumulation test with *Lumbriculus varigatus*.
(Compiled from U.S. EPA 2000)

A. It is recommended that the following performance criteria be met when conducting a 28-d test with *Lumbriculus varigatus*:

1. Numbers of *L. varigatus* in a 4-d toxicity screening should not be significantly reduced in the test sediment relative to the control sediment.

2. Test organisms should burrow into test sediment. Avoidance of test sediment by *L. varigatus* may decrease bioaccumulation

3. Hardness, alkalinity, and ammonia in overlying water typically should not vary by more than 50% during the test, and dissolved oxygen in the overlying water should be maintained above 2.5 mg/L.

B. Performance-based criteria for culturing *L. varigatus* include the following:

1. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicant tests to assess the sensitivity of culture organisms. Data from these reference-toxicant tests could be used to assess genetic strain or life-stage sensitivity of test organisms to selected chemicals.

2. Laboratories should monitor the frequency with which the population is doubling in the culture (number of organisms) and record this information using control charts (doubling rate would need to be estimated on a subset of animals from a mass culture). Records should also be kept on the frequency of restarting cultures. If static cultures are used, it may be desirable to measure water quality more frequently.

3. Food used to culture organisms should be analyzed before the start of a test for compounds to be evaluated in the bioaccumulation test.

4. Laboratories should record the following water-quality characteristics of the cultures at least quarterly and the day before the start of a sediment test: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in cultures should be measured weekly. Temperature of cultures should be recorded daily.

5. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.

Table 34 (continued). Test acceptability requirements for a 28-d sediment bioaccumulation test with *Lumbriculus varigatus*.
(Compiled from U.S. EPA 2000)

---

6. Physiological measurements such as lipid content might provide useful information regarding the health of cultures.

D. Additional requirements:

1. All organisms in a test must be from the same source.

2. Sediment collected from the field should be stored in the dark at 4ºC and should be held for as little time as possible, though actual recommended storage times vary.

3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.

4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.

5. Test organisms must be cultured and tested at 23°C (±1°C).

6. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.

7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

Lumbriculus varigatus Sediment Toxicity Test Strengths, Limitations and Potential Confounding Factors

Compiled from U.S. EPA 2000

The oligochaete *Lumbriculus varigatus* is an appropriate species for bioaccumulation testing for a number of reasons. Oligochaetes are infaunal detritivorous worms that are easily cultured and handled and provide sufficient tissue biomass for residue testing. *L. varigatus* is particularly well-suited for bioaccumulation testing because it is tolerant of sediment physico-chemical characteristics, is exposed to contaminants via both particle ingestion and porewater, and is insensitive to contaminants relative to the other commonly used test species. *L. varigatus* may therefore be exposed to contaminated sediments for sufficiently long periods (*e.g.*, > 28-d) to allow steady state chemical accumulations. In addition, this species does not biotransform PAHs. The response of *L. varigatus* in laboratory bioaccumulation studies has been confirmed with natural populations of oligochaetes. Because *L. varigatus* is not sensitive to contaminants, tests with this species are less useful for toxicity assessments.

Freshwater Toxicity Test Strengths, Limitations and Potential Confounding Factors

See also sections on specific test protocols

The primary strength of the protocols discussed above is that these are standardized methods that have for the most part been subjected to rigorous state and national selection criteria prior to their implementation as water and sediment quality assessment tools. All procedures use ecologically relevant species, and these protocols have been demonstrated to be sensitive to a wide variety of toxicants. Repeated inter- and intralaboratory tests with reference toxicants and complex effluent samples have demonstrated that the precision of the majority of these procedures is comparable to analytical chemistry techniques (ASTM 2000a-j, U.S. EPA 1994c, U.S. EPA 2000). Although all of these protocols are considered to be appropriate for water and sediment quality assessments, the discussion of the strengths, limitations, and confounding factors that may affect the various protocols is intended to further guide their application in Ecological Risk Assessments.

Application of Freshwater Toxicity Tests

The majority of freshwater studies reported in the literature include data for the three EPA chronic toxicity tests (*C. dubia, P. promelas, R. subcapitata*). As discussed earlier, Grothe *et al.* (1996) concluded that acute and chronic protocols developed with these freshwater species are technically sound when conducted according to EPA procedures. The conclusions of the participants of the Pellston Workshop summarized in Grothe *et al.* (1996) were that these tests provide useful toxicologic information on the potential for pollution to impact receiving waters, and a number of examples were provided demonstrating the relationship between toxicity test results and ecosystem impacts. De Vlaming and Norberg-King (1999), and de Vlaming *et al.* (2001) also reviewed studies demonstrating the relationship between toxicity tests with these species (laboratory and *in situ*) and ecological affects in systems impacted by point- and non-point source pollution. The preponderance of evidence suggests that single species toxicity tests are powerful tools for investigations of aquatic pollutant impacts. This section of the review provides further guidance for using the protocols described above for hazard assessment and ERAs.

As discussed previously, U.S. EPA (1991a) recommends that three species (*i.e.*, invertebrate, fish, and alga) be used to screen effluents and ambient freshwater samples for toxicity. Therefore, to determine risk associated with chemicals in freshwater systems, samples should be tested using EPA (U.S. EPA 1994c) chronic protocols with *C. dubia*, *P. promelas*, and *R. subcapitata,* at a minimum. *Raphidocelis* is particularly appropriate where phytotoxic chemicals are of concern. Additional tests should be included depending on site-specific considerations (*e.g., D. magna* and additional fish protocols). For example, embryo-larval development tests with fish (*P. promelas, O. latipes, O. mykiss)* may be useful in situations where teratogenic chemicals may be the primary chemicals of concern. In freshwater systems where amphibians are at risk, the developmental assay with *X. laevis* (*i.e.,* FETAX) or alternative species is appropriate. In most cases, alternative endpoints may be included with all of these assays. A number of biomarker endpoints have been incorporated in tests with both fish and invertebrate species. The protocol with *P. promelas* has been adapted for use as a model to determine effects of chemicals that may induce endocrine disruption. Ankley *et al.* (2001) have also

recently reported an adaptation of the test protocol with *P. promelas* using reproductively mature fathead minnows that is designed to assess affects of reproductive toxicants on reproductive fitness and endocrine function. A number of fish species have been used in studies that include multiple biomarker endpoints (*e.g.*, vitellogenin induction, acetylcholinesterase inhibition, histopathologic abnormalities, metallothionein and heatshock protein induction). A number of these endpoints have also been included in the standardized invertebrate toxicity test protocols discussed above (see Huggett *et al.* 1992 for discussion). In addition, a number of alternative endpoints have been developed for tests with daphnids, algae, and bacteria (see Environment Canada 1999 for discussion).

Because chemicals entering aquatic systems often attach to particles and become associated with sediments, risk to benthic systems should also be assessed in freshwater ERAs. This is particularly important in situations where less water-soluble compounds are of concern. In some cases, the most sensitive species may also be those most likely to be exposed. For example, amphipods (e.g., *H. azteca*), isopods, cladocera, and a variety of aquatic insect species have been shown to be among those most sensitive to pyrethroid pesticides (Giddings *et al.* 2001). Because these compounds attach to particles that are deposited as sediment, protocols with benthic species such as the amphipod *Hyalella azteca* or the midge *Chironomus tentans* would be the most appropriate for ERAs where these chemicals are of concern. The protocols developed for *H. azteca* and *C. tentans* are particularly appropriate for assessing effects of sediment-associated contaminants as part of freshwater ERAs because they are sensitive, and use ecologically relevant species that live in direct contact with the sediment. In addition to survival, the protocols developed with *H. azteca* and *C. tentans* include a number of sublethal endpoints, and may be conducted using both solid-phase sediment and porewater exposures. This latter attribute facilitates application of Toxicity Identification Evaluation procedures with these species. As discussed previously, many of the protocols developed for water column exposures have also been used in sediment testing, and species such as *D. magna* have been demonstrated to be useful in both solid-phase and porewater exposures (Burton *et al.* 2001). Many of these species are appropriate for water column and sediment toxicity studies because they interact with both media in nature. Because bioavailability of

sediment-associated contaminants is a critical issue in ERAs concerned with contaminated freshwater sediments, the bioaccumulation test with the oligochaete *L. varigatus* provides a valuable additional tool to the suite of water sediment toxicity tests available to risk assessors.

PART III: INTEGRATED STUDIES

Routes of exposure vary depending on life stage, life history considerations, trophic relationships, and toxicant and pollutant distributions. For example, sediment-associated pollutants may occur in various phases, including dissolved (*i.e.*, in porewater), attached to particles (*e.g.*, clay, minerals, dissolved organic mater, colloids), complexed with inorganic and organic matrices, and within other living organisms. Exposure to pollutants depends on how an organism interacts with these various compartments and may sometimes occur via multiple routes. Assessments of ecological risk requires an understanding of the complexities of contaminant distributions and the various possible exposure pathways to organisms, as well as factors that may influence contaminant bioavailability and the ecological consequences of exposure and toxicity.

Recent workshops concerned with the application of water column and sediment toxicity tests in environmental assessments have concluded that at a minimum, integrated studies combining multiple measures are necessary to improve our ability to make informed decisions regarding ecological risk (Ingersoll *et al.* 1995a, Grothe *et al.* 1996). Where resources allow, toxicity studies should use multiple species and protocols representing a variety of phyla, endpoints, and exposure pathways. La Point *et al.* (1996) recommended that in addition to chemical analyses, field assessments using water column toxicity tests should also include biological assessments to account for some of the limitations of the Whole Effluent Toxicity (WET) tests to assess bioaccumulation, sediment toxicity, and indirect biotic effects. In assessments of ecological risk associated with contaminated sediments, Ingersoll *et al.* (1995a) also concluded the most appropriate approach is to use multiple test protocols and endpoints and to combine these with chemical measures, bioassessments, and measures of bioaccumulation. The underlying principal in weight-of-evidence or "triad" studies (Chapman *et al.* 1987) is

that multiple lines of evidence reduce uncertainty in ecotoxicologic assessments. Although this document has emphasized laboratory toxicity tests, many of the protocols described above have been adapted for *in situ* water column and sediment exposures, especially for freshwater environmental assessments. *In situ* exposures provide a powerful additional tool for assessing potential impact because they address many of the limitations of laboratory exposures including: continuous exposures to account for transient toxicity due to temporally variable pollution; loss of volatile compounds during sample collection and processing (Burton *et al.* in press); and interactive affects of multiple stressors (*e.g.*, phototoxicity due to PAHs, effects of interaction of contaminants with temperature, flow, turbidity).

Because of the complexities of multiple exposure pathways and factors influencing contaminant bioavailability, there is increasing emphasis on including bioaccumulation and tissue measurements as an additional line of evidence in aquatic toxicology studies (Borgman *et al.* 1991, Chapman *et al.* 1997, Luoma 1996). Tissue concentrations provide a more direct indication of actual dose in traditional bioassay tests (Landrum and Robbins 1990, Borgman *et al.* 1991, Lotufo 1998). Tissue chemistry is particularly useful when used in investigations designed to identify causes of toxicity due to complex chemical mixtures (*e.g.*, Maltby *et al.* 1995). This information is also necessary for developing higher trophic level exposure pathways in web models.

Toxicity Identification Evaluations (TIEs) are an important additional tool for use in the ERA process because they provide investigative means for identifying the direct causes of toxicity. Understanding which chemicals cause toxicity reduces uncertainty in risk assessment and allows resources to be devoted to the primary chemicals of concern. TIE procedures are relatively well developed for both freshwater and marine water column and sediment toxicity assessments (*e.g.*, Ankley *et al.* 1991a; U.S. EPA 1991b, 1993b, 1993c, 1996; Burgess *et al.* 2000). Burton *et al.* (2002) have recommended a tiered approach that incorporates all of these lines of evidence (*i.e.*, laboratory and *in situ* toxicity tests, TIEs, chemical analyses of tissue and exposure media, analyses of physico-chemical properties of test matrices, and site physical characterizations).

This review includes brief descriptions of the various standardized test methods used for aquatic toxicology studies and discusses the strengths and limitations of these

procedures as they apply to ecological risk assessments. Because these procedures are evolving and new techniques are under continued development (*e.g.*, Wood *et al.* 2000, Overmeyer *et al.* 2000, Greve *et al.* 1999), readers of this document are encouraged to regularly consult the relevant literature and recent recommendations from state and federal agencies responsible for implementation of ecological risk assessment procedures in environmental regulation.

REFERENCES

Aller RC. 1988. Benthic fauna and biogeochemical processes in marine sediments. In: Blackburn TH, and Sorenson J, eds. *Nitrogen cycling in coastal marine environments*. John Wiley and sons, New York. P 301-338.

Anderson SL, Harrison FL, Chan G, Moore DH. 1990. Comparison of cellular and whole-animal bioassays for estimation of radiation effects in the polychaete worm *Neanthes arenaceodentata* (Polychaeta). Arch Environ Contam Toxicol 19: 164.

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, Reeve M. 1994. Marine Bioassay Project Seventh Report: Refinement of effluent toxicity testing protocols for four marine species. State Water Resources Control Board Report No. 94-2WQ. Sacramento, CA. 124 pp.

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

Anderson BS, Hunt JW, Hester M, Phillips BM. 1996. Assessment of sediment toxicity at the sediment-water interface. In, G.K. Ostrander (ed.) Techniques in Aquatic Toxicology. Ann Arbor, MI, USA: Lewis Publishers.

Anderson BS, Hunt JW, Tudor S, Newman J, Tjeerdema RS, Fairey R, Oakden J, Bretz C, Wilson CJ, LaCaro F, Stephenson MD, Puckett HM, Long ER, Fleming T, Summers K. 1997. Chemistry, toxicity and benthic community conditions in sediments of selected southern California bays and estuaries. Final Report, State Water Resources Control Board, Sacramento, CA. 140 pp.

Anderson BS, Phillips BM, Hunt JW, Stephenson MD, Puckett HM, Reeve M, Palmer FH. 1998a. Marine Bioassay Project – Ninth Report: Investigations of receiving water toxicity in coastal waters. State Water Resources Control Board, Sacramento, CA.

Anderson BS, Hunt JW, Phillips BM, Tudor S, Fairey R, Newman J, Puckett HM, Stephenson M, Long ER, Wilson CJ, Tjeerdema RS. 1998b. Comparison of marine sediment toxicity test protocols for the amphipod *Rhepoxynius abronius* and the polychaete worm *Nereis (Neanthes) arenaceodentata.* Environ. Toxicol. Chem. 17(5):859-866.

Anderson B, Hunt J, Phillips B, Tjeerdema R, Stoelting M, Fairey R. 1999. History of a Hot Spot – Moss Landing Harbor, California. Conference proceedings, Society of Environmental Toxicology and Chemistry (SETAC), Philadelphia, November, 1999.

Anderson BS, Phillips, BM, Hunt, and Sericano, J. 2000. Investigations of chemicals associated with amphipod mortality at two regional monitoring program stations. Technical report to the San Francisco Estuary Institute – Regional Monitoring Program. San Francisco Estuary Instituite, Richmond, CA.

Anderson BS, Hunt JW, Phillips BM, Fairey RJ, Oakden JM, Puckett HM, Stephenson M, Tjeerdema RS, Long ER, Wilson CJ, Lyons M. 2001a. Sediment Quality in Los Angeles Harbor: A triad assessment. Environ. Toxicol. Chem. 20(2):359-370.

Anderson BS, Hunt JW, Phillips BM, Fairey R, Newman J, Puckett HM, Stephenson M, Taberski KT, Tjeerdema RS. 2001b. Influence of sample manipulation on contaminant flux and toxicity at the sediment-water interface. Mar. Environ. Res. 51:191-211.

Anderson BS, de Vlaming V, Larsen K, Deanovic L, Birosik S, Smith DJ, Hunt JW, Tjeerdema RS. 2002. Causes of toxicity in the Calleguas Creek watershed of southern California. . Environ Monit Assess. 78: 131-151.

Anderson SL, Hose JE, Knezovich JP. 1994. Genotoxic and developmental endpoints effects in sea urchins are sensitive indicators of effects of genotoxic chemicals. Environ. Toxicol. Chem. 13:1033-1041.

Ankley GT, Katko A, Arthur JW. 1990. Identification of ammonia as an important sediment-associated toxicant in the lower Fox River and Green Bay, Wisconsin. Environ. Toxicol. Chem. 9:313-322.

Ankley GT, Schubauer-Berigan MK, Dierkes JR, and Lukasewycz MT. 1991a. Sediment toxicity identification evaluation: Phase I, Phase II, and Phase III. U.S. Environmental Protection Agency. Draft Report 08-91. National Effluent Toxicity Assessment Center.

Ankley GT, Schubauer-Berigan MK, Dierkes JR. 1991b. Predicting the toxicity of bulk sediment to aquatic organisms with aqueous test fractions: porewater vs. elutriate. Environ. Toxicol. Chem. 10:1359-1366.

Ankley GT, Di Toro DM, Hansen DJ, Berry WJ. 1996. Technical basis and proposal for deriving sediment quality criteria for metals. Environ. Toxicol. Chem. 15:2056-2066.

Ankley GT, Jensen KM, Kahl MD, Korte JJ, Makynen EA. 2001. Description and evaluation of a short-term reproduction test with the fathead minnow (*Pimephales promelas*). Environ. Toxicol. Chem. 20:1276-1290.

ASTM. 1996. Standard guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. In, Annual Book of ASTM Standards, Vol 11.05. E 1367-92. West Conshohocken, PA, USA: ASTM. pp. 769-794.

ASTM. 2000a. Standard guide for conducting static and flow-through acute toxicity tests with mysids from the west coast of the United States. In, Annual Book of ASTM Standards, Vol 11.05. E 1463-94. West Conshohocken, PA, USA: ASTM. pp. 828-854.

ASTM. 2000b. In, Annual Book of ASTM Standards, Vol 11.05. E 1022-94. West Conshohocken, PA, USA: ASTM. pp.

ASTM. 2000c. Standard guide for conducting bioconcentration tests with fishes and saltwater bivalve mollusks. In, Annual Book of ASTM Standards, Vol 11.05. E 1022-94. West Conshohocken, PA, USA: ASTM. pp. 320-337.

ASTM. 2000d. Standard guide for determination of the bioaccumulation of sediment-associated contaminants by benthic invertebrates. In, Annual Book of ASTM Standards, Vol 11.05. E 1688-00. West Conshohocken, PA, USA: ASTM. pp. 1059-1112.

ASTM. 2000e. Standard guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. In, Annual Book of ASTM Standards, Vol 11.05. E 1367-99. West Conshohocken, PA, USA: ASTM. pp. 711-738.

ASTM. 2000f. Standard guide for conducting acute, chronic, and life-cycle aquatic toxicity tests with polychaetous annelids. In, Annual Book of ASTM Standards, Vol 11.05. E 1562-94. West Conshohocken, PA, USA: ASTM. pp. 924-943.

ASTM. 2000g. Standard guide for acute toxicity test with the rotifer Branchionus. In, Annual Book of ASTM Standards, Vol 11.05. E 1440-91. West Conshohocken, PA, USA: ASTM. pp. 820-827.

ASTM. 2000h. Standard guide for conducting *Daphnia magna* life-cycle toxicity tests. In, Annual Book of ASTM Standards, Vol 11.05. E 1193-97. West Conshohocken, PA, USA: ASTM. pp. 453-470.

ASTM. 2000i Standard guide for conducting early life-stage toxicity tests with fishes. In, Annual Book of ASTM Standards, Vol 11.05. E 1241-98. West Conshohocken, PA, USA: ASTM. pp. 528-555.

ASTM. 2000j. Standard guide for conducting the frog embryo teratogenesis assay-*Xenopus* (FETAX). In, Annual Book of ASTM Standards, Vol 11.05. E 1439-98. West Conshohocken, PA, USA: ASTM. pp. 804-819.

Bay S, Burgess R, Nacci D. 1993. Status and applications of echinoid (Phylum Echinodermata) toxicity test methods. "*Environmental Toxicology and Risk Assessment*", ASTM STP 1179, Wayne G. Landis, Jane S. Hughes, and Michael A. Lewis, Eds., Philadelphia, PA, USA:American Society for Testing and Materials. pp. 281-302.

Bay S, Jones BH, Schiff K. 1999. Study of the impact of stormwater discharge on Santa Monica Bay. Technical publication USCSG-TR-02-99. Sea Grant program, Wrigley

Institute of Environmental Studies, University of Southern California, Los Angeles, CA. Nov. 1, 1999. 16 pp.

Bay SM, Brown JS, Greenstein DJ, Jirik AW. 2001. Toxicity of methyl-tert-butyl ether (MTBE) to California marine life. In, Weisberg S (ed): Southern California Coastal Water Research Project Annual Report 1999-2000. pp. 136-142.

Bay SM, Anderson BS, Carr RS. 2003. Relative performance of porewater and solid-phase toxicity tests: Characteristics, causes, and consequences. In, Carr RS, Nipper M (eds): *Porewater toxicity testing: biological chemical and ecological considerations – methods, applications, and recommendations for future areas of research*. Pensacola, FL, USA: SETAC Press. pp. 11-36.

Beatty Jr. TV, McFee WE, Fleming CA, Giffin M. 2000. Comparative study of toxicity testing and macroinvertebrate assessment of petroleum storage and manufacturing site and receiving water. Environ. Toxicol. Chem. 19:405-416.

Bedard D, Hayton A, Persaud D. 1992. Ontario Ministry of the Environment Laboratory Sediment Biological Testing Protocol. Ontario Ministry of the Environment, Toronto, Ontario. Pp 26.

Borgman U, Norwood WP, Babirad IM. 1991. Relationship between chronic toxicity and bioaccumulation of cadmium in *Hyalella azteca*. Can. J. Fish. Aquat. Sci. 48:1055-1060.

Bridges TS, Farrar JD. 1997. The influence of worm age, duration of exposure and endpoint selection on bioassay sensitivity for *Neanthes arenaceodentata*. Environ. Toxicol. Chem. 16:1650-1658.

Bridges TS, Farrar JD, Duke BM. 1997. The influence of food ration on sediment toxicity in *Neanthes arenaceodetata*. Environ. Toxicol. Chem. 16:1659-1665.

Bufflap SE, Allen HE. 1995. Sediment pore water collection methods for trace metal analysis: a review. Wat. Res. 29:165-177.

Burgess RM, Cantwell MG, Pelletier MC, Ho KT, Serbst JR, Cook HF, Kuhn A. 2000. Development of toxicity identification evaluation procedure for characterizing metal toxicity in marine sediments. Environ. Toxicol. Chem. 19:981-991.

Burgess RM, Charles JB, Kuhn A, Ho KT, Patton LE, McGovern DG. 1997. Development of a cation-exchange methodology for marine toxicity identification evaluation applications. Environ. Toxicol. Chem. 16:1203-1211.

Burton Jr. GA. 1991. Assessing the toxicity of freshwater sediments. Environ. Toxicol. Chem. 10:1585-1627.

Burton Jr. GA, Pitt R, Clark S. 2000. The role of traditional and novel toxicity test methods in assessing stormwater and sediment contamination. Crit. Rev. Environm. Sci. Technol. 30:413-447.

Burton GA Jr, Batley GE, Chapman PM, Forbes VE, Smith EP, Reynoldson T, Schlekat CE, den Besten PJ, Bailer AJ, Green AS, Dwyer RL. 2002. A weight-of-evidence framework for assessing sediment (or other) contamination: Improving certainty in the decision-making process. Human and Ecolog Risk Assess 8: 1675-1696.

Cairns MA, Nebeker JH, Gakstatter JH, Griffis W. 1984. Toxicity of copper-spiked sediments to freshwater invertebrates. Environ. Toxicol. Chem. 3:435-446.

Canfield TJ, Kemble NE, Brumbaugh WG, Dwyer FJ, Ingersoll CG, Fairfield JF. 1994. Use of benthic invertebrate community structure and the sediment quality triad to evaluate metal-contaminated sediment in the upper Clark Fork River, Montana. Environ. Toxicol. Chem. 13:1999–2012.

Carr RS, Nipper M (eds). 2003. *Porewater toxicity testing: biological chemical and ecological considerations – methods, applications, and recommendations for future areas of research*. Setac Press, Pensacola, FL, USA.

Carr RS, Montagna PA, Biedenbach JM, Kalke R, Kennicutt MC, Hooten R, Cripe G. 2000. Impact of storm-water outfalls on sediment quality in Corpus Christi Bay, Texas, USA. Environ. Toxicol. Chem. 19:561-574.

Chapman PM, Dexter RN, Long ER. 1987. Synoptic measures of sediment contamination, toxicity, and infaunal community structure (the Sediment Quality Triad). Mar. Ecol. Prog. Series 37:75-96.

Chapman PM, Anderson BS, Carr RS, Engle V, Haverland P, Hameedi J, Hyland J, Ingersoll C, Long E, Rodgers J, Sibley S, Smith P, Swartz R, Thompson B, Windom H. 1997. General guidelines for using the Sediment Quality Triad (SQT). Mar. Poll. Bull. 34(6):368-372.

Day KE, Scott IM. 1990. Use of acetylchloinesterase activity to detect sublethal toxicity in stream invertebrates exposed to low concentrations of organophosphate pesticides. Aquatic. Toxicol. 18:101-114.

Day KE, Dutka BJ, Kwan KK, Batista N, Reynoldson TB, Metcalfe-Smith JL. 1995. Correlations between solid-phase microbial screening assays, whole-sediment toxicity tests with macroinvertebrates and *in situ* benthic community structure. J. Great Lakes Res. 21:192-206.

De Vlaming V, Norberg-King TJ. 1999. A review of single species toxicity tests: Are the tests reliable predictors of aquatic ecosystem community responses? EPA 600/R-97/11. Technical report. U.S. Environmental Protection Agency, Duluth, MN.

De Vlaming V, Connor V, DiGiorgio C, Bailey HC, Deanovic LA, Hinton DE. 2000. Application of whole effluent toxicity test procedures to ambient water quality assessment. Environ. Toxicol. Chem. 19:42-62.

De Vlaming, V, Denton D, Crane M. 2001. Multiple lines of evidence are useful, but individual lines of evidence should not be minimized – a reply to Hall and Giddings (2000). Human Ecologic. Risk Assess. 7:443-457.

DeWitt TH, Ditsworth GR, Swartz RC. 1988. Effects of natural sediment features on survival of the phoxocephalid amphipod, *Rhepoxynius abronius*. Mar. Environ. Res. 25:99-124.

DeWitt TH, Swartz RC, Lamberson JO. 1989. Measuring the toxicity of estuarine sediment. Environ. Toxicol. Chem. 8:1035-1048.

DeWitt TH, Glover SA, Emlen JM. 1997. Using multi-generation exposure experiments with amphipods to test predictions of chronic toxicity tests and population models. Abstract. Society of Environmental Toxicology and Chemistry annual meeting, San Francisco, CA.

Dillon TM, Moore DW, Gibson AB. 1993. Development of a chronic sublethal bioassay for evaluating contaminated sediment with the marine polychaete worm *Nereis (Neanthes) arenaceodentata*. Environ. Toxicol. Chem. 12:589-605.

Dillon TM, Moore DW, Reish DJ. 1995. A 28-day sediment bioassay with the marine polychaete, *Nereis (Neanthes) arenaceodentata*. In: *Environmental Toxicology and Risk Assessment* – Third Volume, ASTM 1218. Hughes JS, Biddinger GR, Mones E, eds. American Society for Testing and Materials, Philadelphia, PA. Pp 201- 215.

Dinnel PA, Stober QA, Link JM, Letourneau MW, Roberts WE, Felton SP, Nakatani RE. 1983. Methodology and validation of a sperm cell toxicity test for testing toxic substances in marine waters. Final Report, FRI-UW-8306, Fisheries Reseach Inst., School of Fisheries, University of Washington, Seattle, WA. 208 pp.

Environment Canada. 1999. Guidance document on application and interpretation of single-species tests in environmental toxicology. EPS 1/RM/34 – December 1999. Ontario, Canada. 203 pp. with appendices.

Fairey R, Roberts C, Jacobi M, Lamerdin S, Clark R, Downing J, Long E, Hunt J, Anderson B, Newman J, Stephenson M, Wilson CJ. 1998. Assessment of sediment toxicity and chemical concentrations in the San Diego Bay region, California. Environ. Toxicol. Chem. 17:1570-1581.

Farrar JD, Lotufo GR, Duke BM, Emery Jr V, Bridges TS. 1999. Comparative sensitivity of 10-d acute and 28-d chronic *Leptocheirus plumulosus* toxicity tests. Proceeding: Society of Environmental Toxicology and Chemistry 20th annual meeting, Philadelphia, PA November 14-18. p 29.

Finlayson BJ, Harrington JM, Fojimura R, Isaacs G. 1991. Identification of methyl parathion toxicity in Colusa Basin drain water. Environ. Toxicol. Chem. 12:291-303.

Fisher TC, Crane M, Callaghan A. 2000. An optimized microtiterplate assay to detect acetylchloinesterase activity in individual *Chironomus riparius* meigen. Environ. Toxicol. Chem. 19:1749-1752.

Forbes TL, Forbes VE, Giessing A, Hansen R, Kure LK. 1998. Relative role of porewater versus ingested sediment in bioavailability of organic contaminants in marine sediments. Environ. Toxicol. Chem. 17:2453-2462.

Fort DJ, Rogers RL, Copley HF, Bruning LA, Stover EL, Helhen JC, Burkhart JG. 1999. Progress toward identifying causes of maldevelopment induced in *Xenopus* by pond water and sediment extracts from Minnesota, USA. Environ. Toxicol. Chem. 18:2316-2324.

Giddings JM, Solomon KR, Maund SJ. 2001. Probabilistic risk assessment of cotton pyrethroids: II. Aquatic mesocosm and field studies. Environ. Toxicol. Chem. 20:660-668.

Giesy JP, Rosiu CJ, Graney RL, Heny MG. 1990. Benthic invertebrate bioassays with toxic sediment and porewater. Environ. Toxicol. Chem. 9:233-248.

Green, AS, Chandler GT, Blood ER. 1993. Aqueous-, pore-water-, and sediment-phase cadmium toxicity relationships for meiobenthic copepod. Environ. Toxicol. Chem. 12:1497-1506.

Green A, Moore D, Farrar D. 1999. Chronic toxicity of 2,4,6-trinitrotoluene to a marine polychaete and an estuarine amphipod. Environ. Toxicol. Chem. 18:1783-1790.

Greve GD, Van der Gest HG, Stuijfzand SC, Engels S, Kraak MHS. 1999. Development and validation of an ecotoxicity test using field-collected eggs of the riverine mayfly *Ephoron virgo*. Proc. Exp. Appl. Entomol. Leiden, Nev, Amsterdam 10:105-110.

Gries TH. 1998. Larval bioassay workshop summary. Workshop proceedings sponsored by the Puget Sound dredged materials management program. Washington Department of Ecology. January, 1998.

Gries TH. 2000. Regulatory use of sediment quality values. Workshop proceedings sponsored by the Southern California Coastal Water Research Project, October, 2000.

Grothe DR, Dickson KL, Reed-Judkins DK. (eds). 1996. Whole effluent toxicity testing: an evaluation of methods and prediction of receiving water impacts. SETAC Pellston Workshop on Whole Effluent Toxicity; 1995 Sep 16-25; Pellston, MI. Pensacola, FL, USA:SETAC Press. 340 pp.

Gully JR, Bottomley JP, Baird RB. 1999. Effects of sporophyll storage on giant kelp *Macrocystis pyrifera* (Agardh) bioassay. Environ. Toxicol. Chem. 18:1474-1481.

Helmstetter MF, Maccubbin AE, Alden RW. 1996. The medaka embryo-larval assay: an in vivo assay for toxicity, teratogenicity, and carcinogenicity. In, Ostrander, GK (ed), Techniques in Aquatic Toxicology. Boca Raton, FL, USA:Lewis Publishers. pp. 93-24.

Hemming JM, Allen HJ, Waller WT, Denslow ND, Amman LP. 2000. Assessment of the potential of a treatment wetland to reduce wastewater estrogenicity and toxicity. Abstract presented at the 21st SETAC annual meeting in Nashville, TN, November 12-16.

Ho KT, McKinney RA, Kuhn A, Pelletier MC, Burgess RM. 1997. Identification of acute toxicants in New Bedford Harbor sediments. Environ. Toxicol. Chem. 16:551-558.

Ho K, Caudle D. 1997. Ion toxicity and produced water. Letter to the Editor. Environ. Toxicol. Chem. 16:1993-1995.

Hoke RA, Gala WR, Drake JB, Giesy JP, Flegler S. 1992. Bicarbonate as a potential confounding factor in cladoceran toxicity assessments of porewater from contaminated sediments. Can. J. Fish. Aquat. Sci. 49:1633-1640.

Hooten RL, Carr RS. 1998. Development and application of a marine sediment pre-water toxicity test using *Ulva fasciata* zoopsores. Environ. Toxicol. Chem. 17:932-940.

Huggett RJ, Kimerle RA, Mehrle Jr. PM, Bergman HL. 1992. Biomarkers: Biochemical, physiological, and histological markers of anthropogenic stress. Boca Raton, FL, USA:Lewis Publishers. 347 pp.

Hunt JW, Anderson BS. 1989. Sublethal effects of zinc and municipal sewage effluents on larvae of the red abalone *Haliotis rufescens*. Mar. Biol. 101(4):545-552.

Hunt JW, Anderson BS, Turpen SL, Coulon AR, Martin M, Palmer FH, Janik JJ. 1991. Marine Bioassay Project Sixth Report: Interlaboratory comparisons and protocol development with four marine species. Report No. 91-21-WQ. State Water Resources Control Board, Sacramento, CA. 204 pp.

Hunt JW, Anderson BS. 1993. From research to routine: A review of toxicity testing with marine molluscs. Environmental Toxicology and Risk Assessment, ASTM STP 1179, Wayne G. Landis, Jane S. Hughes, and Michael A. Lewis, Eds, American Society for Testing and Materials, Philadelphia, PA, USA. pp 320-339.

Hunt JW, Anderson BS, Phillips BM, Tjeerdema RS, Puckett HM, deVlaming V. 1999. Patterns of aquatic toxicity in an agriculturally dominated coastal watershed of California. Agricul. Ecosyst. Environ. 75:75-91.

Hunt JW, Anderson BS, Phillips BM, Tjeerdema R, Taberski KM, Wilson CJ, Puckett HM, Stephenson M, Fairey R, Oakden J. 2001a. A large-scale categorization of sites in San Francisco Bay, USA based on the sediment quality triad, toxicity identification evaluations, and gradient studies. Environ. Toxicol. Chem. 20:1252-1265.

Hunt JW, Anderson BS, Phillips BM, Newman J, Tjeerdema R, Fairey R, Puckett HM, Stephenson M, Smith RW, Wilson CJ, Taberski KM. 2001b. Evaluation and use of sediment toxicity reference sites for statistical comparisons in regional assessments. Environ. Toxicol. Chem. 20:1266-1275.

Ingersoll CG, Dillon T, Biddinger GR (eds). 1995a. *Ecological Risk Assessment of Conatminated Sediments*. Pensacola, FL, USA:SETAC Press. 389 pp.

Ingersoll CG, Ankley GT, Benoit DA, Burton GA, Dwyer FJ, Greer IE, Norberg-King TJ, Winger PV. 1995b. Toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates: A review of methods and applications. Environ. Toxicol. Chem. 14:1885-1894.

Ireland DS, Burton GA, Hess GG. 1996. *In situ* toxicity evaluations of turbidity and photoinduction of polycyclic aromatic hydrocarbons. Environ. Toxicol. Chem. 15:574-581.

Johns DM, Ginn TC, Reish DJ. 1990. Protocol for juvenile N. arenaceodentata sediment bioassay. EPA 68-D8-0085. U.S. Environmental Protection Agency, Seattle, WA.

Kemble NE, Brumbaugh WG, Brunson EL, Dwyer FJ, Ingersoll CG, Monda DP, Woodward DF. 1994. Toxicity of metal-contaminated sediments from the Upper Clark Fork River, MT, to aquatic invertebrates in laboratory exposures. Environ. Toxicol. Chem. 13:1985-1997.

Kemble NE, Ingersoll CG, Kunz JL. 2000. Relative sensitivity of endpoints in long-term water or sediment exposures with the amphipod *Hyalella azteca* and the midge *Chironomus tentans*. Abstract presented at the 21st SETAC annual meeting in Nashville, TN, November 12-16.

Knezovich JP, Steichen DJ, Jelinski JA, Anderson SL. 1996. Sulfide tolerance of four marine species used to evaluate sediment and porewater toxicity. Bull. Environ. Contam. Toxicol. 57:450-457.

Kocan RM. 1996. Fish embryos as *in situ* monitors of aquatic pollution. In, *Techniques in Aquatic Toxicology,* GK Ostrander (ed.). Boca Raton, FL, USA:CRC Press. pp 73-92.

Kohn NP, Word JQ, Niyogi DK, Ross LT, Dillon T, Moore DW. 1994. Acute toxicity of ammonia to four species of marine amphipod. Mar. Environ. Res. 38:1-15.

Korte JJ, Kahl MD, Jensen KM, Pasha MS, Ankley GT. 1998. Development of a ribonuclease protection assay for vitellogenin m RNA in fathead minnows as a screen for estrogen receptor agonists. Abstract presented at the 21st SETAC annual meeting in Charlotte, NC, November 15-19.

Lacey R, Watzin MC, McIntosh AW. 1999. Sediment organic matter content as a confounding factor in toxicity tests with *Chironomus tentans*. Environ. Toxicol. Chem. 18:231-236.

Lamberson J, Swartz R, Ozretich R. 1996. Field validation of acute and chronic marine amphipod sediment toxicity tests with *Grandidierella japonica*. Abstract presented at the 17th SETAC annual meeting in Washington, DC, November 17-21.

Landrum PF, Robbins JA. 1990. Bioavailability of sediment-associated contaminants to benthic invertebrates. In, Baudo, R, Giesy, J, Muntau, H (eds): *Sediment: Chemistry and Toxicity of In-Place Pollutants*. pp. 237-264.

Langdon CJ, Harmon VL, Vance MM, Kreeger KE, Kreeger DA, Chapman GA. 1996. A 7-d toxicity test for marine pollutants using the Pacific mysid *Mysidopsis intii*. 1. Culture and protocol development. Environ. Toxicol. Chem. 15:1815-1823.

La Point TW, Barbour MT, Borton DL, Cherry DS, Clements WH, Diamond JM, Grothe DR, Lewis MA, Reed-Judkins DK, Saalfield GW. 1996. Field Assessments – Discussion Synopsis. In, Grothe DR, Dickson KL, Reed-Judkins DK (eds). *Whole effluent toxicity testing: an evaluation of methods and prediction of receiving water impacts.* SETAC Pellston Workshop on Whole Effluent Toxicity; 1995 Sep 16-25; Pellston, MI. Pensacola, FL, USA:SETAC Press. pp 191–228.

Lasier PJ, Winger PV, Reinert RE. 1997. Toxicity of alkalinity to *Hyalella azteca*. Bull. Environ. Contam. Toxicol. 59:807-814.

Lasier PJ, Winger PV, Bogenrieder KJ. 2000. Toxicity of manganese to *Ceriodaphnia dubia* and *Hyalella azteca*. Arch. Environ. Contam. Toxicol. 38:298-304.

Lee BG, Griscom SB, Lee JS, Choi HJ, Koh CH, Luoma SN, Fisher NS. 2000. Influence of dietary uptake and acid-volatile sulfide on bioavailability of metals to sediment-dwelling organisms. Science 287:282-284.

Long ER, Buchman MF. 1989. An evaluation of candidate measures of biological effects for the National Status and Trends Program. NOAA Technical Memorandum. NOS OMA 45. National Atmospheric and Oceanic Administration, Seattle.

Long ER, MacDonald DD, Smith SL, Calder FD. 1995. Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. Environ. Manag. 19:81-97.

Long ER, Field LJ, MacDonald DD. 1998. Predicting toxicity in marine sediments with numerical sediment quality guidelines. Environ. Toxicol. Chem. 17:714-727.

Long ER, Hameedi J, Robertson A, Aasen S, Dutch M, Ricci C, Welch K, Kammin W, Carr RS, Johnson T, Biedenbach J, Scott KJ, Mueller C, Anderson JW. 1999. Survey of sediment quality in Puget Sound. Year 1 – Northern Puget Sound. Draft Technical Memorandum – NOS ORCA. National Oceanic and Atmospheric Administration. Silver Spring, MD.

Long ER. 2000. Degraded sediment quality in U.S. estuaries: a review of magnitude and ecological implications. Ecol Applications 10:338-349.

Long ER, Hong CB, Severn CG. 2001. Relationships between acute sediment toxicity in laboratory tests and abundance and diversity of benthic infauna in marine sediments: a review. Environ Toxicol Chem 20: 46-60.

Lotufo GR. 1998. Lethal and sublethal toxicity of sediment-associated fluoranthene to benthic copepods: application of the critical-body-residue approach. Aquat. Toxxicol. 44:17-30.

Luoma SN. 1996. The developing framework of marine ecotoxicology: Pollutants as a variable in marine ecosystems? J. Exp. Mar. Biol. Ecol. 200:29-55.

Luoma SN, Ho KT. 1993. The appropriate uses of marine and estuarine sediment bioassays. In: Calow, P. editor. Handbook of Ecotoxicology. Blackwell Scientific, Oxford. P 193-228.

Maltby ML, Boxal, ABA, Forrow DM, Calow P, Betton CI. 1995. The effects of motorway runoff on freshwater ecosystems: 2 identifying major toxicants. Environ. Toxicol. Chem. 11:609-614.

McGee BL, Fisher DJ, Yonkos LT, Ziegler GP, Turley S. 1999. Assessment of sediment contamination, acute toxicity, and population variability of the estuarine amphipod Leptocheirus plumulosus in Baltimore Harbor, Maryland, USA. Environ Toxicol Chem 18: 2151-2160.

Meador JP. 1993. The effect of laboratory holding on the toxicity response of marine infaunal amphipods to cadmium and tributlytin. J. Exp. Mar. Biol. Ecol. 174:227-242.

Mearns AJ, Word JQ. 1982. Forecasting effects of sewage solids on marine benthic communities. In: *Ecological Stress and the New York Bight: Science and Mangement*. Mayer GF (ed). Estuarine Research Federation, Columbia, SC. Pp 495.

Mearns AJ, Swartz RC, Cummins JM, Dinnel PA, Plesh P, Chapman PM. 1996. Inter-laboratory comparison of a sediment toxicity test using the marine amphipod Rhepoxynius abronius. Mar Environ Res. 19: 13-37.

Middaugh DP, Anderson BS. 1993. Utilization of topsmelt, *Atherinops affinis*, in environmental toxicology studies along the Pacific coast of the United States. Rev. Environ. Toxicol. 5:1-49.

Nebeker AV, Cairns MA, Gakstatter JH, Malueg KW, Schuytema GS, Krawczyk DF. 1984. Biological methods for determining the toxicity of contaminated freshwater sediments to invertebrates. Environ Toxicol Chem 3: 617-630.

Newsted JL, Giesy JP. 1987. Predictive model for photoinduced acute toxicity of polycyclic aromatic hydrocarbons to *Daphnia magna* Strauss (Cladocera Crustacea). Environ Toxicol Chem: 445-461.

Nicely PA, Hunt JW, Anderson BS, Palmer FA, Carley S. 2000. Tolerance of several marine toxicity test organisms to ammonia and artificial salts. Proceedings, Society of Environmental Toxicology and Chemistry (SETAC), Philadephia, PA.

Nipper MG, Greenstein DJ, Bay SM. 1989. Short- and long-term sediment toxicity test methods with the amphipod *Grandidierella japonica*. Enviro. Toxicol. Chem. 8:1191-1200.

O'Reilly Wiese SB, MacLeod CL, Lester JN. 1997. Partitioning of metals between dissolved and particulate phases in the salt marshes of Essex and North Norfolk (UK). Environ. Technol. 18:399-408.

Oris JT, Giesy JP. 1987. The photoinduced toxicity of polycyclic aromatic hydrocarbons to larvae of the fathead minnow (*Pimephales promelas*). Chemosphere 16:1396-1404.

Overmeyer JP, Noblet R, Armbrust KL. 2000. Black fly larvae (Diptera: Simuliidae): Toxicity evaluations of insecticides entering suburban watersheds. Abstract presented at the 21st SETAC annual meeting in Nashville, TN, November 12-16.

Pastorak RA, Becker DS. 1990. Comparative sensitivity of sediment toxicity bioassays at three superfund sites in Puget Sound. In, Landis WG, van der Schalie WH (eds), *Aquatic Toxicology and Risk Assessment.* Vol. 13, ASTM STP 1096. American Society for Testing and Materials, Philadelphia, PA, USA. pp. 123-139.

Pesch CE. 1983. *Neanthes arenaceodentata* (Polychaetous annelid), a proposed cytogenetic model for marine toxicology. In: Workshop on sublethal effects of stress on marine organisms. Martin M, Harrison F (eds). Asilomar, Pacific Grove, CA. California State Water Resources Control Board, Sacramento, CA. p 73.

Pesch CE, Hoffman GL. 1983. Interlaboratory comparison of a 28-day toxicity test with the polychaete *Neanthes arenaceodentata*. In: Aquatic toxicology and hazard assessment sixth symposium. Bishop WE et al. (eds). ASTM STP802, American Society for Testing and Materials, Philadelphia, PA. Pp 482-493.

Phillips, BM, Nicely, PA, Hunt, JW, Anderson, BS, Tjeerdema, RS, Palmer, SE, Palmer, FH, Puckett, HM. 2003. Toxicity of cadmium-copper-nickel-zinc mixtures to larval purple sea urchins (*Strongylocentrotus purpuratus*) and bay mussels (*Mytilus galloprovincialis*). Bull. Environ Contam. Toxicol. 70: 592-599.

Phillips, B.M., Anderson, B.S., and Hunt, J.W. 1997. Measurement and distribution of interstitial and overlying water ammonia and hydrogen sulfide in sediment toxicity tests. Mar. Environ. Res. 44(2): 117-126.

Phillips BM, Anderson BS, Hunt JW. 2000. Investigation of causes of sediment toxicity in the San Francisco Bay/Delta. Technical report to the San Francisco Estuary Institute – Regional Monitoring Program, SFEI, Richmond, CA.

Pillard DA, DuFresne DL, Caudle DD, Tietge JE, Evans JM. 2000. Predicting the toxicity of major ions in seawater to mysid shrimp (*Mysidopsis bahia*), sheepshead minnow (*Cyprinodon variegatus*), and inland silverside minnow (*Menidia beryllina*). Environ. Toxicol. Chem. 19:183-191.

Redmond MS, Scott KJ, Swartz RC, Jones JKP. 1994. Preliminary culture and life-cycle experiments with the benthic amphipod *Ampelisca abdita*. Environ Toxicol Chem 13: 1355-1365.

Reish DJ. 1985. The use of the polychaetous annelid *Neanthes arenaceodentata* as a laboratory experimental animal. Tethys 11: 335-341.

Reish DJ, Gerlinger TV. 1984. The effect of cadmium, lead, and zinc in the polychaetous annelid *Neanthes arenaceodentata*. Proc. Linnean Soc. New South Wales. pp. 383-389.

Reynoldson TB, Day KE, Clark C, Milani D. 1994. Effects of indigenous animals on chronic endpoints in freshwater sediment toxicity tests. Environ. Contam. Toxicol. 13:937-977.

Roberts MH. 1987. Acute toxicity of tributyltin chloride to embryos and larvae of two bivalve mollusks, *Crassostrea virginica* and *Mercenaria mercenaria*. Bull. Environ. Contam. Toxicol. 39:1012-1019.

Robson DL. 1990. Observations of dye dispersal in the interstitial and overlying waters with the benthic amphipod Ampelisca abdita. Proceedings of the 11th annual meeting of the Society of Environmental Toxicology and Chemistry, November 11-15, Arlington, VA, p 138.

Schiff KC, Reish DJ, Anderson JW, Bay SM. 1992. A comparative evaluation of produced water toxicity. In, *Produced Water*. Ray, JP, Engelhart, FR (eds). New York:Plenum Press. pp 199-207.

Schimmel SC, Morrison GE, Heber MA. 1989. Marine complex effluent toxicity test program: test sensitivity, repeatability and relevance to receiving water toxicity. Environ. Toxicol. Chem. 8:739-746.

Schimmel SC, Thursby GB. 1996. Predicting receiving system impacts from effluent toxicity: a marine perspective. In, Grothe DR, Dickson KL, Reed-Judkins DK (eds). *Whole effluent toxicity testing: an evaluation of methods and prediction of receiving water impacts.* SETAC Pellston Workshop on Whole Effluent Toxicity; 1995 Sep 16-25; Pellston, MI. Pensacola, FL, USA: SETAC Press. pp. 322-330.

Schlekat CE, Scott KJ, Swartz RC, Albrecht B, Antrim L, Doe K, Douglas S, Ferretti JA, Hansen DJ, Moore DW, Mueller C, Tang A. 1995. Interlaboratory comparison of a 10-day sediment toxicity test method using *Ampelisca abdita, Eohaustorius estuarius*, and *Leptocheirus plumulosus*. Environ. Toxicol. Chem 14:2163-2174.

Schubauer-Berigan MK, Ankley GT. 1991. The contribution of ammonia, metals, and nonpolar organic compounds to the toxicity of sediment interstitial water from an Illinois River tributary. Environ. Toxicol. Chem. 10:925-940.

Scott KJ, Redmond MS. 1989. The effects of a contaminated dredge material on laboratory populations of the tubicolous amphipod, Ampelisca abdita. In: *Aquatic Toxicology and Hazard Assessment* 12th Volume. Cowgill UM, Williams LR (eds). STP 1027, American Society of Testing and Materials, Philadelphia, PA. pp 289-303.

Sibley PK, Legler J, Dixon DG, Barton DR. 1997a. Environmental health assessment of benthic habitat adjacent to a pulp mill discharge. I. Acute and chronic toxicity of sediments to benthic macroinvertebrates. Arch. Environ. Contam. Toxicol. 32:274-284.

Sibley PK, Benoit DA, Ankley GT. 1997b. The significance of growth in *Chironomus tentans* sediment toxicity tests: Relationship to reproduction and demographic endpoints. Environ. Toxicol. Chem. 16:336-345.

Sibley PK, Benoit DA, Balcer MD, Phipps GL, West CW, Hoke RA, Ankley GT. 1999. *In situ* bioassay chamber for assessment of sediment toxiciy and bioaccumulation using benthic invertebrates. Environ. Toxicol. Chem. 18:2325-2336.

Sibley PK, Chappel MJ, George TK, Solomon KR, Liber K. 2000. Integrating effects of stressors across levels of biological organization: examples using organophosphorus insecticide mixtures in field-level exposures. J. Aquat. Ecosyst. Stress Recov. 7:117-130.

Singer MM, George S, Lee I, Jacobson S, Weetman LL, Blondina G, Tjeerdema RS. 1998. Effects of dispersant treatment on the acute toxicity of petroleum hydrocarbons. Arch. Environ. Contam. Toxicol. 34:177-187.

Suter GW, Barnthouse LW, Efroymson RA, Jager H. 1999. Ecological risk assessment in a large river-reservoir: 2. Fish community. Environ. Toxicol. Chem. 18:589-598.

Swartz RC, DeBen WA, Cole FA. 1979. A bioassay for the toxicity of sediment to marine macrobenthos. Journal of the Water Pollution Control Federation 5:944-950.

Swartz RC, Deben WA, Sercu KA, Lamberson JO. 1982. Sediment toxicity and the distribution of amphipods in Commencement Bay, Washington, USA. Mar. Poll. Bull. 13: 359-364.

Swartz RC, DeBen WA, Jones JKP, Lamberson JO, Cole FA. 1985. Phoxocephalid amphipod bioassay for marine sediment toxicity. In, Aquatic toxicity and hazard assessment, Seventh Symposium: ASTM STP 854, R.D. Cradwell. R. Purdy and R.C. Bahner (eds), American Society for Testing and Materials, Philadelphia, PA, USA. pp. 284-307.

Swartz RC, Cole FA, Schults DW, Deben WA. 1986. Ecological changes in the Southern California Bight near a large sewage outfall: benthic conditions in 1980 and 1983. Mar Ecol Prog Ser 31: 1-13.

Swartz RC, Kemp PF, Shults DW, Ditsworth GR, Ozretich RJ. 1989. Acute toxicity of sediment from Eagle Harbor, Washington to the infaunal amphipod, Rhepoxynius abronius. Environ. Toxicol. Chem. 8:215-222.

Swartz RC, Cole FA, Lamberson JO, Ferraro SP, Schults DW, Deben WA, Lee II H, Ozretich RJ. 1994. Sediment toxicity, contamination and amphipod abundance at a DDT- and Dieldrin-contaminated site in San Francisco Bay. Environ. Toxicol. Chem. 13:949-962.

SWRCB. 1996. Procedures manual for conducting toxicity tests developed by the Marine Bioassay Project. Report No. 96-1WQ. State Water Resources Control Board, Sacramento, CA. 627 pp.

SWRCB. 2000. Final functional equivalent document, amendment of the water quality control plan for ocean waters of California: California Ocean Plan. State Water Resources Control Board, Sacramento, CA. pp A24-A26.

Thompson B, Anderson B, Hunt J, Taberski K, Phillips B. 1999. Relationships between sediment toxicity and contamination in San Francisco Bay. Mar. Environ. Res. 48:285-309.

Thursby GB, Anderson BS, Walsh GE, Steele RL. 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, Wayne G. Landis, Jane S. Hughes, and Michael A. Lewis (eds), American Society for Testing and Materials, Philadelphia, PA, USA. pp 362-377.

Tietge JE, Hockett JR, Evans JM. 1997. Major ion toxicity of six produced waters to three freshwater species: application of ion toxicity models and TIE procedures. Environ. Toxicol. Chem. 16:2002-2008.

Tietge JE, Ankley GT, DeFoe DL, Holcombe GW, Jensen KM. 2000. Effects of water quality on development of *Xenopus laevis*: a frog embryo teratogenis assay- *Xenopus* assessment of surface water associated with malformations in native anurans. Environ. Toxicol. Chem. 19:2114-2121.

U.S. EPA. 1991. Technical Support Document for Water Quality-Based Toxics Control. Office of Water. EPA/505/2-90-001.

U.S. EPA. 1991. Methods for Aquatic Toxicity Identification Evaluations: Phase I Toxicity Characterization Procedures. Office of Research and Development. EPA 600-6-91-003. February, 1991.

U.S. EPA. 1993a. Methods for measureing the acute toxicity of effluents and receiving waters to freshwater and marine organisms. Fourth edition, Weber CI (ed). Environmental Monitoring Systems Laboratory, Cincinnati, OH, Office of Research and Development, U.S. Environmental Protection Agency. 293 pp.

U.S. EPA. 1993b. Methods for Aquatic Toxicity Identification Evaluations: Phase II Toxicity Identification Procedures for samples exhibiting acute and chronic toxicity. Office of Research and Development. EPA 600-R-92-080. September, 1993.

U.S. EPA. 1993c. Methods for Aquatic Toxicity Identification Evaluations: Phase III Toxicity Confirmation Procedures for samples exhibiting acute and chronic toxicity. Office of Research and Development. EPA 600-R-92-081. September, 1993.

U.S. EPA. 1994a. Short-term methods for estimating the chronic toxicity of effluents and receiving water to marine and estuarine organisms. Office of Research and Development. EPA 600-4-91-003. June 1994.

U.S. EPA. 1994b. Methods for Assessing the Toxicity of Sediment-associated Contaminants with Estuarine and Marine Amphipods. Office of Research and Development. EPA 600-R-94-025. June 1994.

U.S. EPA. 1994c. Short-term methods for estimating the chornic toxicity of effluents and receiving water to freshwater organisms. Third edition, Lewis PA, Klemm DJ, Lazorchak JM, Norberg-King, TJ, Peltier WH, Heber MA (eds). Environmental Monitoring Systems Laboratory, Cincinnati, OH, Office of Research and Development, U.S. Environmental Protection Agency. 341 pp.

U.S. EPA. 1995a. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to west coast marine and estuarine organisms. EPA/600/R-95/136. pp 321-388.

U.S. EPA. 1995b. Environmental Monitoring and Assessment Program (EMAP): laboratory methods manual – estuaries, Volume 1 – biological and physical analyses. United States Environmental Protection Agency, Office of Research and Development, Narragansett, RI. EPA/620/R-95/008.

U.S. EPA. 1996. Marine Toxicity Identification Evaluation (TIE): Phase I Guidance Document. Office of Research and Development. EPA 600-R-96-054. September, 1996.

U.S. EPA. 1998. Evaluation of dredged material proposed for discharge in waters of the U.S.-inland testing manual. EPA-823-B-98-004. U.S. EPA ACOE.

U.S. EPA. 2000. Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates. Office of Research and Development. EPA 600-R-99-096. March 2000.

Van Ginneken L, Chowdhury MJ, Blust R. 1999. Bioavailability of cadmium and zinc to the common carp, *Cyrpinus carpio*, in complexing environments: a test for the validity of the fee ion activity model. Environ. Toxicol. Chem. 18:2295-2304.

Vismann B. 1990. Sulfide detoxification and tolerance in *Nereis (Hediste) diversicolor* and *Nereis (Neanthes) virens* (Annelida:Polychaeta). Mar. Ecol. Prog. Ser. 59:229-238.

Wang F, Chapman PM. 1999. Biological implications of sulfide in sediment--a review focusing on sediment toxicity. Environ. Toxicol. Chem. 18:2526-2532.

Werner I, Nagel R. 1997. Stress proteins HSP60 and HSP70 in three species of amphipods exposed to cadmium, diazinon, dieldrin, and fluoranthene. Environ. Toxicol. Chem. 16:293-2403.

Weston DP. 1995. Further development of a chronic *Ampelisca abdita* bioassay as an indicator of sediment toxicity: summary and conclusions. Technical Report, Regional Monitoring Program for Trace Substances. San Francisco Estuary Institute, Richmond, CA. pp 108-115.

Winger PV, Lasier PJ. 1991. A vacuum-operated pore-water extractor for estuarine and freshwater sediments. Arch. Environ. Contam. Toxicol. 21:321-324.

Winger PV, Lasier PJ, White DH, Seginal JT. 2000. Effects of contaminants in dredge material from the lower Savannah River. Arch. Environ. Contam. Toxicol. 38:128-136.

Winger PV, Albrecht B, Anderson BS, Bay SM, Bona F, Stephenson GL. In review. Comparison of Porewater and Solid-Phase Sediment Toxicity Tests. In, Carr RS, Nipper M (eds): *Porewater toxicity testing: biological chemical and ecological considerations – methods, applications, and recommendations for future areas of research*. Pensacola, FL, USA:SETAC Press.

Woelke CE. 1967. Measurement of water quality with the Pacific Oyster embryo bioassay. Special Technical Publication 416, American Society for Testing and Materials, Philadelphia, PA, pp 112-120.

Wood RM, Maltby L, Brock TCM, Gylstra R. 2000. Using whole-organism physiological and biochemical responses to assess the toxicity of lambda-cyhalothrin to *Gammarus pulex*. Abstract presented at the 21st SETAC annual meeting in Nashville, TN, November 12-16.

Word JQ, Ward JA, Franklin LM, Cullinan VI, Kiesser SL. 1987. Evaluation of the equilibrium partitioning theory for estimating the toxicity of the nonpolar organic compound DDT to the sediment dwelling amphipod Rhepoxynius abronius. CSD 11, U.S. Environmental Protection Agency, Washington, DC.