Designation: E1367 – 03(Reapproved 2008)

Standard Test Method for Measuring the Toxicity of Sediment-Associated Contaminants with Estuarine and Marine Invertebrates[1]

This standard is issued under the fixed designation E1367; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope*

1.1 This test method covers procedures for testing estuarine or marine organisms in the laboratory to evaluate the toxicity of contaminants associated with whole sediments. Sediments may be collected from the field or spiked with compounds in the laboratory. General guidance is presented in Sections 1-15 for conducting sediment toxicity tests with estuarine or marine amphipods. Specific guidance for conducting 10-d sediment toxicity tests with estuarine or marine amphipods is outlined in Annex A1 and specific guidance for conducting 28-d sediment toxicity tests with *Leptocheirus plumulosus* is outlined in Annex A2.

1.2 Procedures are described for testing estuarine or marine amphipod crustaceans in 10-d laboratory exposures to evaluate the toxicity of contaminants associated with whole sediments (Annex A1; USEPA 1994a (**1**)). Sediments may be collected from the field or spiked with compounds in the laboratory. A toxicity method is outlined for four species of estuarine or marine sediment-burrowing amphipods found within United States coastal waters. The species are *Ampelisca abdita*, a marine species that inhabits marine and mesohaline portions of the Atlantic coast, the Gulf of Mexico, and San Francisco Bay; *Eohaustorius estuarius*, a Pacific coast estuarine species; *Leptocheirus plumulosus*, an Atlantic coast estuarine species; and *Rhepoxynius abronius*, a Pacific coast marine species. Generally, the method described may be applied to all four species, although acclimation procedures and some test conditions (that is, temperature and salinity) will be species-specific (Sections 12 and Annex A1). The toxicity test is conducted in 1-L glass chambers containing 175 mL of sediment and 775 mL of overlying seawater. Exposure is static (that is, water is not renewed), and the animals are not fed over the 10-d exposure period. The endpoint in the toxicity test is survival with reburial of surviving amphipods as an additional measurement that can be used as an endpoint for some of the test species (for *R. abronius* and *E. estuarius*). Performance criteria established for this test include the average survival of amphipods in negative control treatment must be greater than or equal to 90 %. Procedures are described for use with sediments with pore-water salinity ranging from >0 ‰ to fully marine.

1.3 A procedure is also described for determining the chronic toxicity of contaminants associated with whole sediments with the amphipod *Leptocheirus plumulosus* in laboratory exposures (Annex A2; USEPA-USACE 2001(**2**)). The toxicity test is conducted for 28 d in 1-L glass chambers containing 175 mL of sediment and about 775 mL of overlying water. Test temperature is 25° ± 2°C, and the recommended overlying water salinity is 5 ‰ ± 2 ‰ (for test sediment with pore water at 1 ‰ to 10 ‰) or 20 ‰ ± 2 ‰ (for test sediment with pore water >10 ‰). Four hundred millilitres of overlying water is renewed three times per week, at which times test organisms are fed. The endpoints in the toxicity test are survival, growth, and reproduction of amphipods. Performance criteria established for this test include the average survival of amphipods in negative control treatment must be greater than or equal to 80 % and there must be measurable growth and reproduction in all replicates of the negative control treatment. This test is applicable for use with sediments from oligohaline to fully marine environments, with a silt content greater than 5 % and a clay content less than 85 %.

1.4 A salinity of 5 or 20 ‰ is recommended for routine application of 28-d test with *L. plumulosus* (Annex A2; USEPA-USACE 2001 (**2**)) and a salinity of 20 ‰ is recommended for routine application of the 10-d test with *E. estuarius* or *L. plumulosus* (Annex A1). However, the salinity of the overlying water for tests with these two species can be adjusted to a specific salinity of interest (for example, salinity representative of site of interest or the objective of the study may be to evaluate the influence of salinity on the bioavailability of chemicals in sediment). More importantly, the salinity tested must be within the tolerance range of the test organisms (as outlined in Annex A1 and Annex A2). If tests are conducted with procedures different from those described in 1.3 or in Table A1.1 (for example, different salinity, lighting, temperature, feeding conditions), additional tests are required to determine comparability of results (1.10). If there is not a

[1] This test method is under the jurisdiction of ASTM Committee E50 on Environmental Assessment, Risk Management and Corrective Action and is the direct responsibility of Subcommittee E50.47 on Biological Effects and Environmental Fate.

Current edition approved Feb. 1, 2008. Published April 2008. Originally approved in 1990. Last previous edition approved in 2003 as E1367 – 03ε1 . DOI: 10.1520/E1367-03R08.

*A Summary of Changes section appears at the end of this standard

TABLE 1 Rating of Selection Criteria for Estuarine or Marine Amphipod Sediment Toxicity Testing A "+" or "–" Rating Indicates a Positive or Negative Attribute

Criterion	*Ampelisca abdita*	*Eohaustorius estuarius*	*Leptocheirus plumulosus*	*Rhepoxynius abronius*
Relative sensitivity toxicity data base	+	+	+	+
Round-robin studies conducted	+	+	+	+
Contact with sediment	+	+	+	+
Laboratory culture	+/-	-	+	-
Taxonomic identification	+	+	+	+
Ecological importance	+	+	+	+
Geographical distribution	ATL, PAC, GOM	PAC	ATL	PAC
Sediment physicochemical tolerance	+	+	+	+
Response confirmed with benthos populations	+	+[A]	+	+
Peer reviewed	+	+	+	+
Endpoints monitored	Survival	Survival, reburial	Survival	Survival, reburial

[A] Anderson et al. (2001 (**40**)).

ATL = Atlantic Coast, PAC = Pacific Coast, GOM= Gulf of Mexico

need to make comparisons among studies, then the test could be conducted just at a selected salinity for the sediment of interest.

1.5 Future revisions of this standard may include additional annexes describing whole-sediment toxicity tests with other groups of estuarine or marine invertebrates (for example, information presented in Guide E1611 on sediment testing with polychaetes could be added as an annex to future revisions to this standard). Future editions to this standard may also include methods for conducting the toxicity tests in smaller chambers with less sediment (Ho et al. 2000 (**3**), Ferretti et al. 2002 (**4**)).

1.6 Procedures outlined in this standard are based primarily on procedures described in the USEPA (1994a (**1**)), USEPA-USACE (2001(**2**)), Test Method E1706, and Guides E1391, E1525, E1688, Environment Canada (1992 (**5**)), DeWitt et al. (1992a (**6**); 1997a (**7**)), Emery et al. (1997 (**8**)), and Emery and Moore (1996 (**9**)), Swartz et al. (1985 (**10**)), DeWitt et al. (1989 (**11**)), Scott and Redmond (1989 (**12**)), and Schlekat et al. (1992 (**13**)).

1.7 Additional sediment toxicity research and methods development are now in progress to *(1)* refine sediment spiking procedures, *(2)* refine sediment dilution procedures, *(3)* refine sediment Toxicity Identification Evaluation (TIE) procedures, *(4)* produce additional data on confirmation of responses in laboratory tests with natural populations of benthic organisms (that is, field validation studies), and *(5)* evaluate relative sensitivity of endpoints measured in 10- and 28-d toxicity tests using estuarine or marine amphipods. This information will be described in future editions of this standard.

1.8 Although standard procedures are described in Annex A2 of this standard for conducting chronic sediment tests with *L. plumulosus*, further investigation of certain issues could aid in the interpretation of test results. Some of these issues include further investigation to evaluate the relative toxicological sensitivity of the lethal and sublethal endpoints to a wide variety of chemicals spiked in sediment and to mixtures of chemicals in sediments from contamination gradients in the field (USEPA-USACE 2001 (**2**)). Additional research is needed to evaluate the ability of the lethal and sublethal endpoints to estimate the responses of populations and communities of benthic invertebrates to contaminated sediments. Research is also needed to link the toxicity test endpoints to a field-validated population model of *L. plumulosus* that would then generate estimates of population-level responses of the amphipod to test sediments and thereby provide additional ecologically relevant interpretive guidance for the laboratory toxicity test.

1.9 This standard outlines specific test methods for evaluating the toxicity of sediments with *A. abdita*, *E. estuarius*, *L. plumulosus*, and *R. abronius*. While standard procedures are described in this standard, further investigation of certain issues could aid in the interpretation of test results. Some of these issues include the effect of shipping on organism sensitivity, additional performance criteria for organism health, sensitivity of various populations of the same test species, and confirmation of responses in laboratory tests with natural benthos populations.

1.10 General procedures described in this standard might be useful for conducting tests with other estuarine or marine organisms (for example, *Corophium spp.*, *Grandidierella japonica*, *Lepidactylus dytiscus*, *Streblospio benedicti*), although modifications may be necessary. Results of tests, even those with the same species, using procedures different from those described in the test method may not be comparable and using these different procedures may alter bioavailability. Comparison of results obtained using modified versions of these procedures might provide useful information concerning new concepts and procedures for conducting sediment tests with aquatic organisms. If tests are conducted with procedures different from those described in this test method, additional tests are required to determine comparability of results. General procedures described in this test method might be useful for conducting tests with other aquatic organisms; however, modifications may be necessary.

1.11 *Selection of Toxicity Testing Organisms:*

1.11.1 The choice of a test organism has a major influence on the relevance, success, and interpretation of a test. Furthermore, no one organism is best suited for all sediments. The following criteria were considered when selecting test organisms to be described in this standard (Table 1 and Guide E1525). Ideally, a test organism should: *(1)* have a toxicological database demonstrating relative sensitivity to a range of

contaminants of interest in sediment, *(2)* have a database for interlaboratory comparisons of procedures (for example, round-robin studies), *(3)* be in direct contact with sediment, *(4)* be readily available from culture or through field collection, *(5)* be easily maintained in the laboratory, *(6)* be easily identified, *(7)* be ecologically or economically important, *(8)* have a broad geographical distribution, be indigenous (either present or historical) to the site being evaluated, or have a niche similar to organisms of concern (for example, similar feeding guild or behavior to the indigenous organisms), *(9)* be tolerant of a broad range of sediment physico-chemical characteristics (for example, grain size), and *(10)* be compatible with selected exposure methods and endpoints (Guide E1525). Methods utilizing selected organisms should also be *(11)* peer reviewed (for example, journal articles) and *(12)* confirmed with responses with natural populations of benthic organisms.

1.11.2 Of these criteria (Table 1), a database demonstrating relative sensitivity to contaminants, contact with sediment, ease of culture in the laboratory or availability for field-collection, ease of handling in the laboratory, tolerance to varying sediment physico-chemical characteristics, and confirmation with responses with natural benthic populations were the primary criteria used for selecting *A. abdita*, *E. estuarius*, *L. plumulosus*, and *R. abronius* for the current edition of this standard for 10-d sediment tests (Annex A1). The species chosen for this method are intimately associated with sediment, due to their tube- dwelling or free-burrowing, and sediment ingesting nature. Amphipods have been used extensively to test the toxicity of marine, estuarine, and freshwater sediments (Swartz et al., 1985 (**10**); DeWitt et al., 1989 (**11**); Scott and Redmond, 1989 (**12**); DeWitt et al., 1992a (**6**); Schlekat et al., 1992 (**13**)). The selection of test species for this standard followed the consensus of experts in the field of sediment toxicology who participated in a workshop entitled "Testing Issues for Freshwater and Marine Sediments". The workshop was sponsored by USEPA Office of Water, Office of Science and Technology, and Office of Research and Development, and was held in Washington, D.C. from 16-18 September 1992 (USEPA, 1992 (**14**)). Of the candidate species discussed at the workshop, *A. abdita*, *E. estuarius*, *L. plumulosus*, and *R. abronius* best fulfilled the selection criteria, and presented the availability of a combination of one estuarine and one marine species each for both the Atlantic (the estuarine *L. plumulosus* and the marine *A. abdita*) and Pacific (the estuarine *E. estuarius* and the marine *R. abronius*) coasts. *Ampelisca abdita* is also native to portions of the Gulf of Mexico and San Francisco Bay. Many other organisms that might be appropriate for sediment testing do not now meet these selection criteria because little emphasis has been placed on developing standardized testing procedures for benthic organisms. For example, a fifth species, *Grandidierella japonica* was not selected because workshop participants felt that the use of this species was not sufficiently broad to warrant standardization of the method. Environment Canada (1992 (**5**)) has recommended the use of the following amphipod species for sediment toxicity testing: *Amphiporeia virginiana*, *Corophium volutator*, *Eohaustorius washingtonianus*, *Foxiphalus xiximeus*, and *Leptocheirus pinguis*. A database similar to those available for *A. abdita*, *E. estuarius*, *L. plumulosus*, and *R. abronius* must be developed in order for these and other organisms to be included in future editions of this standard.

1.11.3 The primary criterion used for selecting *L. plumulosus* for chronic testing of sediments was that this species is found in both oligohaline and mesohaline regions of estuaries on the East Coast of the United States and is tolerant to a wide range of sediment grain size distribution (USEPA-USACE 2001 (**2**), Annex Annex A2). This species is easily cultured in the laboratory and has a relatively short generation time (that is, about 24 d at 23°C, DeWitt et al. 1992a(**6**)) that makes this species adaptable to chronic testing (Section 12).

1.11.4 An important consideration in the selection of specific species for test method development is the existence of information concerning relative sensitivity of the organisms both to single chemicals and complex mixtures. Several studies have evaluated the sensitivities of *A. abdita*, *E. estuarius*, *L. plumulosus*, or *R. abronius*, either relative to one another, or to other commonly tested estuarine or marine species. For example, the sensitivity of marine amphipods was compared to other species that were used in generating saltwater Water Quality Criteria. Seven amphipod genera, including *Ampelisca abdita* and *Rhepoxynius abronius*, were among the test species used to generate saltwater Water Quality Criteria for 12 chemicals. Acute amphipod toxicity data from 4-d water-only tests for each of the 12 chemicals was compared to data for *(1)* all other species, *(2)* other benthic species, and *(3)* other infaunal species. Amphipods were generally of median sensitivity for each comparison. The average percentile rank of amphipods among all species tested was 57 %; among all benthic species, 56 %; and, among all infaunal species, 54 %. Thus, amphipods are not uniquely sensitive relative to all species, benthic species, or even infaunal species (USEPA 1994a (**1**)). Additional research may be warranted to develop tests using species that are consistently more sensitive than amphipods, thereby offering protection to less sensitive groups.

1.11.5 Williams et al. (1986 (**15**)) compared the sensitivity of the *R. abronius* 10-d whole sediment test, the oyster embryo (*Crassostrea gigas*) 48-h abnormality test, and the bacterium (*Vibrio fisheri*) 1-h luminescence inhibition test (that is, the Microtox[2] test) to sediments collected from 46 contaminated sites in Commencement Bay, WA. *Rhepoxynius abronius* were exposed to whole sediment, while the oyster and bacterium tests were conducted with sediment elutriates and extracts, respectfully. Microtox[2] was the most sensitive test, with 63 % of the sites eliciting significant inhibition of luminescence. Significant mortality of *R. abronius* was observed in 40 % of test sediments, and oyster abnormality occurred in 35 % of sediment elutriates. Complete concordance (that is, sediments that were either toxic or not-toxic in all three tests) was observed in 41 % of the sediments. Possible sources for the lack of concordance at other sites include interspecific differences in sensitivity among test organisms, heterogeneity in contaminant types associated with test sediments, and differences in routes of exposure inherent in each toxicity test. These

[2] Microtox is a trademark of Strategic Diagnostics Inc. 111 Pencader Drive Newark, Delaware 19702-3322.

results highlight the importance of using multiple assays when performing sediment assessments.

1.11.6 Several studies have compared the sensitivity of combinations of the four amphipods to sediment contaminants. For example, there are several comparisons between *A. abdita* and *R. abronius*, between *E. estuarius* and *R. abronius*, and between *A. abdita* and *L. plumulosus*. There are fewer examples of direct comparisons between *E. estuarius* and *L. plumulosus*, and no examples comparing *L. plumulosus* and *R. abronius*. There is some overlap in relative sensitivity from comparison to comparison within each species combination, which appears to indicate that all four species are within the same range of relative sensitivity to contaminated sediments.

1.11.6.1 Word et al. (1989 (**16**)) compared the sensitivity of *A. abdita* and *R. abronius* to contaminated sediments in a series of experiments. Both species were tested at 15°C. Experiments were designed to compare the response of the organism rather than to provide a comparison of the sensitivity of the methods (that is, *Ampelisca abdita* would normally be tested at 20°C). Sediments collected from Oakland Harbor, CA, were used for the comparisons. Twenty-six sediments were tested in one comparison, while 5 were tested in the other. Analysis of results using Kruskal Wallace rank sum test for both experiments demonstrated that *R. abronius* exhibited greater sensitivity to the sediments than *A. abdita* at 15°C. Long and Buchman (1989 (**17**)) also compared the sensitivity of *A. abdita* and *R. abronius* to sediments from Oakland Harbor, CA. They also determined that *A. abdita* showed less sensitivity than *R. abronius*, but they also showed that *A. abdita* was less sensitive to sediment grain size factors than *R. abronius*.

1.11.6.2 DeWitt et al. (1989 (**11**)) compared the sensitivity of *E. estuarius* and *R. abronius* to sediment spiked with fluoranthene and field-collected sediment from industrial waterways in Puget Sound, WA, in 10-d tests, and to aqueous cadmium ($CdCl_2$) in a 4-d water-only test. The sensitivity of *E. estuarius* was from two (to spiked-spiked sediment) to seven (to one Puget Sound, WA, sediment) times less sensitive than *R. abronius* in sediment tests, and ten times less sensitive to $CdCl_2$ in the water-only test. These results are supported by the findings of Pastorok and Becker (1990 (**18**)) who found the acute sensitivity of *E. estuarius* and *R. abronius* to be generally comparable to each other, and both were more sensitive than *Neanthes arenaceodentata* (survival and biomass endpoints), *Panope generosa* (survival), and *Dendraster excentricus* (survival).

1.11.6.3 *Leptocheirus plumulosus* was as sensitive as the freshwater amphipod *Hyalella azteca* to an artificially created gradient of sediment contamination when the latter was acclimated to oligohaline salinity (that is, 6 ‰ ; McGee et al., 1993 (**19**)). DeWitt et al. (1992b (**20**)) compared the sensitivity of *L. plumulosus* with three other amphipod species, two mollusks, and one polychaete to highly contaminated sediment collected from Baltimore Harbor, MD, that was serially diluted with clean sediment. *Leptocheirus plumulosus* was more sensitive than the amphipods *Hyalella azteca* and *Lepidactylus dytiscus* and exhibited equal sensitivity with *E. estuarius*. Schlekat et al. (1995 (**21**)) describe the results of an interlaboratory comparison of 10-d tests with *A. abdita*, *L. plumulosus* and *E. estuarius* using dilutions of sediments collected from Black Rock Harbor, CT. There was strong agreement among species and laboratories in the ranking of sediment toxicity and the ability to discriminate between toxic and non-toxic sediments.

1.11.6.4 Hartwell et al. (2000 (**22**)) evaluated the response of *Leptocheirus plumulosus* (10-d survival or growth) to the response of the amphipod *Lepidactylus dytiscus* (10-d survival or growth), the polychaete *Streblospio benedicti* (10-d survival or growth), and lettuce germination (Lactuca sativa in 3-d exposure) and observed that *L. plumulosus* was relatively insensitive compared to the response of either L. dytiscus or S. benedicti in exposures to 4 sediments with elevated metal concentrations.

1.11.6.5 Ammonia is a naturally occurring compound in marine sediment that results from the degradation of organic debris. Interstitial ammonia concentrations in test sediment can range from <1 mg/L to in excess of 400 mg/L (Word et al., 1997 (**23**)). Some benthic infauna show toxicity to ammonia at concentrations of about 20 mg/L (Kohn et al., 1994 (**24**)). Based on water-only and spiked-sediment experiments with ammonia, threshold limits for test initiation and termination have been established for the *L. plumulosus* chronic test. Smaller (younger) individuals are more sensitive to ammonia than larger (older) individuals (DeWitt et al., 1997a(**7**), b (**25**). Results of a 28-d test indicated that neonates can tolerate very high levels of pore-water ammonia (>300 mg/L total ammonia) for short periods of time with no apparent long-term effects (Moore et al., 1997 (**26**)). It is not surprising *L. plumulosus* has a high tolerance for ammonia given that these amphipods are often found in organic rich sediments in which diagenesis can result in elevated pore-water ammonia concentrations. Insensitivity to ammonia by *L. plumulosus* should not be construed as an indicator of the sensitivity of the *L. plumulosus* sediment toxicity test to other chemicals of concern.

1.11.7 Limited comparative data is available for concurrent water-only exposures of all four species in single-chemical tests. Studies that do exist generally show that no one species is consistently the most sensitive.

1.11.7.1 The relative sensitivity of the four amphipod species to ammonia was determined in ten-d water only toxicity tests in order to aid interpretation of results of tests on sediments where this toxicant is present (USEPA 1994a (**1**)). These tests were static exposures that were generally conducted under conditions (for example, salinity, photoperiod) similar to those used for standard 10-d sediment tests. Departures from standard conditions included the absence of sediment and a test temperature of 20°C for *L. plumulosus*, rather than 25°C as dictated in this standard. Sensitivity to total ammonia increased with increasing pH for all four species. The rank sensitivity was *R. abronius* = *A. abdita* > *E. estuarius* > *L. plumulosus*. A similar study by Kohn et al. (1994 (**24**)) showed a similar but slightly different relative sensitivity to ammonia with *A. abdita* > *R. abronius* = *L. plumulosus* > *E. estuarius*.

1.11.7.2 Cadmium chloride has been a common reference toxicant for all four species in 4-d exposures. DeWitt et al. (1992a (**6**)) reports the rank sensitivity as *R. abronius* > *A. abdita* > *L. plumulosus* > *E. estuarius* at a common temperature and salinity of 15°C and 28 ‰ . A series of 4-d exposures

to cadmium that were conducted at species-specific temperatures and salinities showed the following rank sensitivity: *A. abdita* = *L. plumulosus* = *R. abronius* > *E. estuarius* (USEPA 1994a (**1**)).

1.11.7.3 Relative species sensitivity frequently varies among contaminants; consequently, a battery of tests including organisms representing different trophic levels may be needed to assess sediment quality (Craig, 1984 (**27**); Williams et al. 1986 (**15**); Long et al., 1990 (**28**); Ingersoll et al., 1990 (**29**); Burton and Ingersoll, 1994 (**31**)). For example, Reish (1988 (**32**)) reported the relative toxicity of six metals (arsenic, cadmium, chromium, copper, mercury, and zinc) to crustaceans, polychaetes, pelecypods, and fishes and concluded that no one species or group of test organisms was the most sensitive to all of the metals.

1.11.8 The sensitivity of an organism is related to route of exposure and biochemical response to contaminants. Sediment-dwelling organisms can receive exposure from three primary sources: interstitial water, sediment particles, and overlying water. Food type, feeding rate, assimilation efficiency, and clearance rate will control the dose of contaminants from sediment. Benthic invertebrates often selectively consume different particle sizes (Harkey et al. 1994 (**33**)) or particles with higher organic carbon concentrations which may have higher contaminant concentrations. Grazers and other collector-gatherers that feed on aufwuchs and detritus may receive most of their body burden directly from materials attached to sediment or from actual sediment ingestion. In some amphipods (Landrum, 1989 (**34**)) and clams (Boese et al., 1990 (**35**)) uptake through the gut can exceed uptake across the gills for certain hydrophobic compounds. Organisms in direct contact with sediment may also accumulate contaminants by direct adsorption to the body wall or by absorption through the integument (Knezovich et al. 1987 (**36**)).

1.11.9 Despite the potential complexities in estimating the dose that an animal receives from sediment, the toxicity and bioaccumulation of many contaminants in sediment such as Kepone®, fluoranthene, organochlorines, and metals have been correlated with either the concentration of these chemicals in interstitial water or in the case of non-ionic organic chemicals, concentrations in sediment on an organic carbon normalized basis (Di Toro et al. 1990 (**37**); Di Toro et al. 1991 (**38**)). The relative importance of whole sediment and interstitial water routes of exposure depends on the test organism and the specific contaminant (Knezovich et al. 1987 (**36**)). Because benthic communities contain a diversity of organisms, many combinations of exposure routes may be important. Therefore, behavior and feeding habits of a test organism can influence its ability to accumulate contaminants from sediment and should be considered when selecting test organisms for sediment testing.

1.11.10 The use of *A. abdita*, *E. estuarius*, *R. abronius*, and *L. plumulosus* in laboratory toxicity studies has been field validated with natural populations of benthic organisms (Swartz et al. 1994 (**39**) and Anderson et al. 2001 (**40**) for *E. estuarius*, Swartz et al. 1982 (**43**) and Anderson et al. 2001 (**40**) for *R. abronius*, McGee et al. 1999 (**41**) and McGee and Fisher 1999 (**42**) for *L. plumulosus*).

1.11.10.1 Data from USEPA Office of Research and Development's Environmental Monitoring and Assessment program were examined to evaluate the relationship between survival of *Ampelisca abdita* in sediment toxicity tests and the presence of amphipods, particularly ampeliscids, in field samples. Over 200 sediment samples from two years of sampling in the Virginian Province (Cape Cod, MA, to Cape Henry, VA) were available for comparing synchronous measurements of *A. abdita* survival in toxicity tests to benthic community enumeration. Although species of this genus were among the more frequently occurring taxa in these samples, ampeliscids were totally absent from stations that exhibited *A. abdita* test survival <60 % of that in control samples. Additionally, ampeliscids were found in very low densities at stations with amphipod test survival between 60 and 80 % (USEPA 1994a (**1**)). These data indicate that tests with this species are predictive of contaminant effects on sensitive species under natural conditions.

1.11.10.2 Swartz et al. (1982 (**43**)) compared sensitivity of *R. abronius* to sediment collected from sites in Commencement Bay, WA, to benthic community structure at each site. Mortality of *R. abronius* was negatively correlated with amphipod density, and phoxocephalid amphipods were ubiquitously absent from the most contaminated areas.

1.11.10.3 Sediment toxicity to amphipods in 10-d toxicity tests, field contamination, and field abundance of benthic amphipods were examined along a sediment contamination gradient of DDT (Swartz et al. 1994 (**39**)). Survival of *E. estuarius* and *R. abronius* in laboratory toxicity tests was positively correlated to abundance of amphipods in the field and along with the survival of *H. azteca*, was negatively correlated to DDT concentrations. The threshold for 10-d sediment toxicity in laboratory studies was about 300 ug DDT (+metabolites)/g organic carbon. The threshold for abundance of amphipods in the field was about 100 ug DDT (+metabolites)/g organic carbon. Therefore, correlations between toxicity, contamination, and biology indicate that acute 10-d sediment toxicity tests can provide reliable evidence of biologically adverse sediment contamination in the field.

1.11.10.4 As part of a comprehensive sediment quality assessment in Baltimore Harbor, MD, McGee et al. (1999 (**41**)) conducted 10-d toxicity tests with *L. plumulosus*. Negative relationships were detected between amphipod survival and concentrations of select sediment-associated contaminants, whereas a very strong positive association existed between survival in laboratory exposures and field density of *L. plumulosus* at test sites. A field validation study of the 10- and 28-d *L. plumulosus* tests by McGee and Fisher (1999 (**42**)) in Baltimore Harbor, also indicated good agreement between acute toxicity, sediment associated contaminants and responses of the *in situ* benthic community. In this study, the chronic 28-d test was less sensitive to sediment contamination than the acute 10-d test; however, the feeding regime used in this evaluation is different than the one currently recommended in Annex A2 and may have influenced the test results. Field validation studies with the revised 28-d test outlined in Annex A2 have not been conducted.

1.12 *Chronic Sediment Methods with Leptocheirus plumulosus:*

1.12.1 Most standard whole sediment toxicity tests have been developed to produce a lethality endpoint (survival/mortality) with potential for a sublethal endpoint (reburial) in some species (USEPA 1994a (**1**), USEPA-USACE 2001 (**2**)). Methods that measure sublethal effects have not been available or have not been routinely used to evaluate sediment toxicity in marine or estuarine sediments (Scott and Redmond, 1989 (**12**); Green and Chandler, 1996 (**44**); Levin et al., 1996 (**45**); Ciarelli et al., 1998 (**46**); Meador and Rice, 2001 (**47**)). Most assessments of contaminated sediment rely on short-term lethality tests (for example, ≤10 d; USEPA-USACE, 1991 (**48**); 1998 (**49**)). Short-term lethality tests are useful in identifying "hot spots" of sediment contamination, but might not be sensitive enough to evaluate moderately contaminated areas. However, sediment quality assessments using sublethal responses of benthic organisms, such as effects on growth and reproduction, have been used to successfully evaluate moderately contaminated areas (Ingersoll et al., 1998 (**50**); Kemble et al., 1994 (**51**); McGee et al., 1995 (**52**); Scott, 1989 (**53**)). The 28-d toxicity test with *Leptocheirus plumulosus* has two sublethal endpoints: growth and reproduction. These sublethal endpoints have potential to exhibit a toxic response from chemicals that otherwise might not cause acute effects or significant mortality in a test. Sublethal response to chronic exposure is also valuable for population modeling of contaminant effects. These data can be used for population-level risk assessments of benthic pollutant effects.

1.12.2 An evaluation of the distribution of *L. plumulosus* in Chesapeake Bay indicates that its distribution is negatively correlated with the degree of sediment contamination (Pfitzenmeyer, 1975 (**54**); Reinharz, 1981 (**55**)). A field validation study of the 10- and 28-d *L. plumulosus* tests by McGee and Fisher (1999 (**42**)) in Baltimore Harbor, indicated good agreement between acute toxicity, sediment associated contaminants and responses of the *in situ* benthic community. In this study, the chronic 28-d test was less sensitive to sediment contamination than the acute 10-d test and therefore had a poorer association between sediment contaminants and benthic community health. It should be noted that the feeding regime used in this evaluation is different than the one currently recommended in Annex A2 and may have influenced the test results. Field validation studies with the revised 28-d test have not been conducted.

1.13 *Limitations*—While some safety considerations are included in this standard, it is beyond the scope of this standard to encompass all safety requirements necessary to conduct sediment tests.

1.14 This standard is arranged as follows:

1.15 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*Specific hazard statements are given in Section 8.

2. Referenced Documents

2.1 *ASTM Standards:*[3]

D1129 Terminology Relating to Water

D4447 Guide for Disposal of Laboratory Chemicals and Samples

E29 Practice for Using Significant Digits in Test Data to Determine Conformance with Specifications

E105 Practice for Probability Sampling of Materials

E122 Practice for Calculating Sample Size to Estimate, With Specified Precision, the Average for a Characteristic of a Lot or Process

E141 Practice for Acceptance of Evidence Based on the Results of Probability Sampling

E177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods

E178 Practice for Dealing With Outlying Observations

E456 Terminology Relating to Quality and Statistics

E691 Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method

E729 Guide for Conducting Acute Toxicity Tests on Test Materials with Fishes, Macroinvertebrates, and Amphibians

E943 Terminology Relating to Biological Effects and Environmental Fate

E1241 Guide for Conducting Early Life-Stage Toxicity Tests with Fishes

E1325 Terminology Relating to Design of Experiments

E1391 Guide for Collection, Storage, Characterization, and Manipulation of Sediments for Toxicological Testing and for Selection of Samplers Used to Collect Benthic Invertebrates

E1402 Guide for Sampling Design

E1525 Guide for Designing Biological Tests with Sediments

[3] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

E1611 Guide for Conducting Sediment Toxicity Tests with Polychaetous Annelids
E1688 Guide for Determination of the Bioaccumulation of Sediment-Associated Contaminants by Benthic Invertebrates
E1706 Test Method for Measuring the Toxicity of Sediment-Associated Contaminants with Freshwater Invertebrates
E1847 Practice for Statistical Analysis of Toxicity Tests Conducted Under ASTM Guidelines
E1850 Guide for Selection of Resident Species as Test Organisms for Aquatic and Sediment Toxicity Tests
IEEE/ASTM SI 10 American National Standard for Use of the International System of Units (SI): The Modern Metric System

3. Terminology

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

3.2 *Definitions*—For definitions of other terms used in this test method, refer to Guides E729 and E1241 and Terminology E943 and D1129. For an explanation of units and symbols, refer to IEEE/ASTM SI 10IEEE/ASTM SI 10.

3.3 *Definitions of Terms Specific to This Standard:*

3.3.1 *clean, n*—denotes a sediment or water that does not contain concentrations of test materials which cause apparent stress to the test organisms or reduce their survival.

3.3.2 *concentration, n*—the ratio of weight or volume of test material(s) to the weight or volume of sediment.

3.3.3 *contaminated sediment, n*—sediment containing chemical substances at concentrations that pose a known or suspected threat to environmental or human health.

3.3.4 *control sediment, n*—a sediment that is essentially free of contaminants and is used routinely to assess the acceptability of a test. Any contaminants in control sediment may originate from the global spread of pollutants and does not reflect any substantial input from local or non-point sources. Comparing test sediments to control sediments is a measure of the toxicity of a test sediment beyond inevitable background contamination.

3.3.5 *EC50, n*—a statistically or graphically estimated concentration that is expected to cause one or more specified effects in 50 % of a group of organisms under specified conditions.

3.3.6 *formulated sediment, n*—mixtures of materials used to mimic the physical components of a natural sediment.

3.3.7 *IC50, n*—a point estimate of the toxicant concentration that would cause a 50 % reduction in a non-quantal measurement such as fecundity or growth.

3.3.8 *interstitial water or pore water, n*— water occupying space between sediment or soil particles.

3.3.9 *LC50, n*—a statistically or graphically estimated concentration that is expected to be lethal to 50 % of a group of organisms under specified conditions.

3.3.10 *lowest-observable-effect concentration (LOEC), n*—in a toxicity test, the lowest tested concentration of a material at which organisms were adversely affected compared to control organisms as determined by statistical hypothesis tests-should be accompanied by a description of the statistical tests and alternative hypotheses, levels of significance, and measures of performance, for example, survival, growth, reproduction, or development-and must be above any other concentration not producing statistically significant adverse effects.

3.3.11 *no-observable-effect concentration (NOEC), n—in a toxicity test*, the highest tested concentration of a material at which organisms did as well as control organisms as determined by statistical hypothesis tests-should be accompanied by a description of the statistical tests and alternative hypotheses, levels of significance, and measures of performance, for example, survival, growth, reproduction, or development-and must be below any other concentration producing statistically significant adverse effects.

3.3.12 *overlying water, n*—the water placed over sediment in a test chamber during a test.

3.3.13 *reference sediment, n*—a whole sediment near an area of concern used to assess sediment conditions exclusive of material(s) of interest. The reference sediment may be used as an indicator of localized sediment conditions exclusive of the specific pollutant input of concern. Such sediment would be collected near the site of concern and would represent the background conditions resulting from any localized pollutant inputs as well as global pollutant input. This is the manner in which reference sediment is used in dredge material evaluations.

3.3.14 *reference-toxicity test, n*—a test conducted with reagent-grade reference chemical to assess the sensitivity of the test organisms. Deviations outside an established normal range may indicate a change in the sensitivity of the test organism population. Reference-toxicity tests are most often performed in the absence of sediment.

3.3.15 *sediment, n*—particulate material that usually lies below water. Formulated particulate material that is intended to lie below water in a test.

3.3.16 *spiked sediment, n*—a sediment to which a material has been added for experimental purposes.

3.3.17 *whole sediment, n*—sediment and associated pore water which have had minimal manipulation. The term bulk sediment has been used synonymously with whole sediment.

4. Summary of Standard

4.1 *Method Description*—Procedures are described for testing estuarine or marine amphipod crustaceans in the 10-d laboratory exposures to evaluate the toxicity of contaminants associated with whole sediments. Sediments may be collected from the field or spiked with compounds in the laboratory. A toxicity method is outlined for four species of estuarine or marine sediment-burrowing amphipods found within United States coastal waters. The species are *Ampelisca abdita*, a marine species that inhabits marine and mesohaline portions of the Atlantic coast, the Gulf of Mexico, and San Francisco Bay; *Eohaustorius estuarius*, a Pacific coast estuarine species; *Leptocheirus plumulosus*, an Atlantic coast estuarine species; and *Rhepoxynius abronius*, a Pacific coast marine species. Generally, the method described may be applied to all four species, although acclimation procedures and some test conditions (that is, temperature and salinity) will be species-specific (Sections 10 and 11). The toxicity test is conducted in 1-L glass chambers containing 175 mL of sediment and 775 mL of overlying seawater. Exposure is static (that is, water is not renewed), and the animals are not fed over the 10-d exposure period. The endpoint in the toxicity test is survival with reburial of surviving amphipods as an additional measurement that can be used as an endpoint for some of the test species (for *R. abronius* and *E. estuarius*). Performance criteria established for this test include the average survival of amphipods in negative control treatment must be greater than or equal to 90 %. Procedures are described for use with sediments with pore-water salinity ranging from >0 % to fully marine.

4.2 A procedure is also described for determining the chronic toxicity of contaminants associated with whole sediments with the amphipod *Leptocheirus plumulosus* in laboratory exposures (USEPA-USACE 2001 (**2**)). The toxicity test is conducted for 28 d in 1-L glass chambers containing 175 mL of sediment and about 775 mL of overlying seawater. Four hundred millilitres of overlying water is renewed three times per week, at which time test organisms are fed. Tests are initiated with neonate amphipods that mature and reproduce during the 28-d test period. The endpoints in the 28-d toxicity test are survival, growth rate, and reproduction of amphipods. Survival is calculated as the percentage of newly born (neonate) amphipods at test initiation that survive as adults at test termination. Growth rate is calculated as the mean dry weight gain per day per adult amphipod surviving at test termination. Reproduction is calculated as the number of offspring per surviving adult. This test is applicable for use with sediment having pore-water salinity ranging from 1 ‰ to 35 ‰ . Typically, endpoint selection for new toxicity tests is generally guided by methodologies for related toxicity tests (Gray et al., 1998 (**56**)). Sediment toxicity tests using macroinvertebrates often incorporate survival and growth endpoints (Ingersoll, 1995 (**57**)). Gray et al. (1998 (**56**)) recommend optimal endpoint measures for the *L. plumulosus* sediment toxicity test based on four criteria: relevance of each measure to its respective endpoint; signal-to-noise ratio (the ratio between the response to stressor and the normal variation in the response variable); redundancy to other measures of the same endpoint; and cost of labor, training, and equipment. Signal-to-noise ratios are independent of experiment design considerations (that is, Type I and Type II errors, and sample size) and are positively correlated with power (Gray et al., 1998 (**56**)).

4.3 *Experimental Design*—The following section is a general summary of experimental design. See Section 13 for additional detail.

4.3.1 *Control and Reference Sediment:*

4.3.1.1 Sediment tests include a control sediment (sometimes called a negative control). A control sediment is a sediment that is essentially free of contaminants and is used routinely to assess the acceptability of a test and is not necessarily collected near the site of concern. Any contaminants in control sediment are thought to originate from the global spread of pollutants and do not reflect any substantial inputs from local or non-point sources Ankley and Thomas, 1992 (**58**). Comparing test sediments to control sediments is a measure of the toxicity of a test sediment beyond inevitable background contamination and organism health Ankley and Thomas, 1992 (**58**). A control sediment provides a measure of test acceptability, evidence of test organism health, and a basis for interpreting data obtained from the test sediments. A reference sediment is collected near an area of concern and is used to assess sediment conditions exclusive of material(s) of interest. Testing a reference sediment provides a site-specific basis for evaluating toxicity.

4.3.1.2 In general, the performance of test organisms in the negative control is used to judge the acceptability of a test, and either the negative control or reference sediment may be used to evaluate performance in the experimental treatments, depending on the purpose of the study. Any study in which organisms in the negative control do not meet performance criteria must be considered questionable because it suggests that adverse factors affected the response of test organisms. Key to avoiding this situation is using only control sediments that have a demonstrated record of performance using the same test procedure. This includes testing of new collections from sediment sources that have previously provided suitable control sediment.

4.3.1.3 Because of the uncertainties introduced by poor performance in the negative control, such studies should be repeated to insure accurate results. However, the scope or sampling associated with some studies may make it difficult or impossible to repeat a study. Some researchers have reported cases where performance in the negative control is poor, but performance criteria are met in reference sediment included in the study design. In these cases, it might be reasonable to infer that other samples that show good performance are probably not toxic; however, any samples showing poor performance should not be judged to have shown toxicity, since it is unknown whether the adverse factors that caused poor control performance might have also caused poor performance in the test treatments.

4.3.1.4 Natural physico-chemical characteristics such as sediment texture may influence the response of test organisms

(59). The physico-chemical characteristics of test sediment need to be within the tolerance limits of the test organism. Ideally, the limits of a test organism should be determined in advance; however, controls for factors including grain size and organic carbon can be evaluated if the limits are exceeded in a test sediment. See section 12.1 and Annex A1 and Annex A2 for information on physico-chemical requirements of test organisms. If the physico-chemical characteristics of a test sediment exceed the tolerance range of the test organism, a control sediment encompassing these characteristics can be evaluated. The effects of sediment characteristics on the results of sediment tests can be addressed with regression equations Dewitt et al. 1988, (59), Ankley et al., 1994(60). The use of formulated sediment can also be used to evaluate physico-chemical characteristics of sediment on test organisms Walsh et al., 1991 (61) Suedel and Rodgers, 1994, (64) Kembel et al.,(63) USEPA, 2000,(62), section 7.2 and Guide E1391).

4.3.2 The experimental design depends on the purpose of the study. Variables that need to be considered include the number and type of control sediments, the number of treatments and replicates, and water quality characteristics. For instance, the purpose of the study might be to determine a specific endpoint such as an LC50 and may include a control sediment, a positive control, a solvent control, and several concentrations of sediment spiked with a chemical (see section 10.3.2). A useful summary of field sampling design is presented by Green, 1979 (65). See Section 13 for additional guidance on experimental design and statistics.

4.3.2.1 The purpose of the study might be to determine if field-collected sediments are toxic and may include controls, reference sediments, and test sediments. Controls are used to evaluate the acceptability of the test (Table A1.3 in Annex A1 and Table A2.3 in Annex A2) and might include a control sediment or a formulated sediment (section 7.2). Testing a reference sediment provides a site-specific basis for evaluating toxicity of the test sediments. Comparisons of test sediments to multiple reference or control sediments representative of the physical characteristics of the test sediment (that is, grain size, organic carbon) may be useful in these evaluations. A summary of field sampling design is presented by Green, 1979 (65). See Section 13 for additional guidance on experimental design and statistics.

4.3.2.2 If the purpose of the study is to conduct a reconnaissance field survey to identify sites for further investigation, the experimental design might include only one sample from each site to allow for sampling a larger area. The lack of replication at a site usually precludes statistical comparisons (for example, analysis of variance (ANOVA)), but these surveys can be used to identify sites for further study or may be evaluated using regression techniques.

4.3.2.3 In other instances, the purpose of the study might be to conduct a quantitative sediment survey of chemistry and toxicity to determine statistically significant differences between effects among control and test sediments from several sites. The number of replicates/site should be based on the need for sensitivity or power (see Section 13). In a quantitative survey, field replicates (separate samples from different grabs collected at the same site) would need to be taken at each site. Chemical and physical characterizations of each of these grabs would be required for each of these field replicates used in sediment testing. Separate subsamples might be used to determine within-sample variability or for comparisons of test procedures (for example, comparative sensitivity among test organisms), but these subsamples cannot be considered to be true field replicates for statistical comparisons among sites.

4.3.2.4 Sediments often exhibit high spatial and temporal variability (66). Therefore, replicate samples may need to be collected to determine variance in sediment characteristics. Sediment should be collected with as little disruption as possible; however, subsampling, compositing, or homogenization of sediment samples may be required for some experimental designs.

4.3.2.5 Site locations might be distributed along a known pollution gradient, in relation to the boundary of a disposal site, or at sites identified as being contaminated in a reconnaissance survey. Comparisons can be made in both space and time. In pre-dredging studies, a sampling design can be prepared to assess the contamination of samples representative of the project area to be dredged. Such a design may include compositing cores collected to project depth from a specified dredged material management area.

4.3.2.6 The primary focus of the physical and experimental test design and statistical analysis of the data, is the experimental unit, which is defined as the smallest physical entity to which treatments can be independently assigned (Guide E1241). Because overlying water or air cannot flow from one test chamber to another the test chamber is the experimental unit. The experimental unit is defined as the smallest physical entity to which treatments can be independently assigned and to which air and water exchange between test chambers are kept to a minimum. Because of factors that might affect results within test chambers and results of a test, all test chambers should be treated as similarly as possible. Treatments should be randomly assigned to individual test chamber locations. Assignment of test organisms to test chambers should be impartial (Guide E729). As the number of test chambers/treatment increases, the number of degrees of freedom increases, and, therefore, the width of the confidence interval on a point estimate, such as an LC50, decreases, and the power of a significance test increases (see Section 13).

5. Significance and Use

5.1 *General:*

5.1.1 Sediment provides habitat for many aquatic organisms and is a major repository for many of the more persistent chemicals that are introduced into surface waters. In the aquatic environment, most anthropogenic chemicals and waste materials including toxic organic and inorganic chemicals eventually accumulate in sediment. Mounting evidences exists of environmental degradation in areas where USEPA Water Quality Criteria (WQC; Stephan et al.(67)) are not exceeded, yet organisms in or near sediments are adversely affected Chapman, 1989 (68). The WQC were developed to protect organisms in the water column and were not directed toward protecting organisms in sediment. Concentrations of contaminants in sediment may be several orders of magnitude higher

than in the overlying water; however, whole sediment concentrations have not been strongly correlated to bioavailability Burton, 1991 (**69**). Partitioning or sorption of a compound between water and sediment may depend on many factors including: aqueous solubility, pH, redox, affinity for sediment organic carbon and dissolved organic carbon, grain size of the sediment, sediment mineral constituents (oxides of iron, manganese, and aluminum), and the quantity of acid volatile sulfides in sediment Di Toro et al. 1991(**70**) Giesy et al. 1988 (**71**). Although certain chemicals are highly sorbed to sediment, these compounds may still be available to the biota. Chemicals in sediments may be directly toxic to aquatic life or can be a source of chemicals for bioaccumulation in the food chain.

5.1.2 The objective of a sediment test is to determine whether chemicals in sediment are harmful to or are bioaccumulated by benthic organisms. The tests can be used to measure interactive toxic effects of complex chemical mixtures in sediment. Furthermore, knowledge of specific pathways of interactions among sediments and test organisms is not necessary to conduct the tests Kemp et al. 1988, (**72**). Sediment tests can be used to: *(1)* determine the relationship between toxic effects and bioavailability, *(2)* investigate interactions among chemicals, *(3)* compare the sensitivities of different organisms, *(4)* determine spatial and temporal distribution of contamination, *(5)* evaluate hazards of dredged material, *(6)* measure toxicity as part of product licensing or safety testing, *(7)* rank areas for clean up, and *(8)* estimate the effectiveness of remediation or management practices.

5.1.3 A variety of methods have been developed for assessing the toxicity of chemicals in sediments using amphipods, midges, polychaetes, oligochaetes, mayflies, or cladocerans (Test Method E1706, Guide E1525, Guide E1850; Annex A1, Annex A2; USEPA, 2000 (**73**), EPA 1994b, (**74**), Environment Canada 1997a, (**75**), Enviroment Canada 1997b,(**76**)). Several endpoints are suggested in these methods to measure potential effects of contaminants in sediment including survival, growth, behavior, or reproduction; however, survival of test organisms in 10-day exposures is the endpoint most commonly reported. These short-term exposures that only measure effects on survival can be used to identify high levels of contamination in sediments, but may not be able to identify moderate levels of contamination in sediments (USEPA USEPA, 2000 (**73**); Sibley et al.1996, (**77**); Sibley et al.1997a, (**78**); Sibley et al.1997b, (**79**); Benoit et al.1997, (**80**); Ingersoll et al.1998, (**81**)). Sublethal endpoints in sediment tests might also prove to be better estimates of responses of benthic communities to contaminants in the field, Kembel et al. 1994 (**82**). Insufficient information is available to determine if the long-term test conducted with *Leptocheirus plumulosus* (Annex A2) is more sensitive than 10-d toxicity tests conducted with this or other species.

5.1.3.1 The decision to conduct short-term or long-term toxicity tests depends on the goal of the assessment. In some instances, sufficient information may be gained by measuring sublethal endpoints in 10-day tests. In other instances, the 10-day tests could be used to screen samples for toxicity before long-term tests are conducted. While the long-term tests are needed to determine direct effects on reproduction, measurement of growth in these toxicity tests may serve as an indirect estimate of reproductive effects of contaminants associated with sediments (Annex A1).

5.1.3.2 Use of sublethal endpoints for assessment of contaminant risk is not unique to toxicity testing with sediments. Numerous regulatory programs require the use of sublethal endpoints in the decision-making process (Pittinger and Adams, 1997, (**83**)) including: *(1)* Water Quality Criteria (and State Standards); *(2)* National Pollution Discharge Elimination System (NPDES) effluent monitoring (including chemical-specific limits and sublethal endpoints in toxicity tests); *(3)* Federal Insecticide, Rodenticide and Fungicide Act (FIFRA) and the Toxic Substances Control Act (TSCA, tiered assessment includes several sublethal endpoints with fish and aquatic invertebrates); *(4)* Superfund (Comprehensive Environmental Responses, Compensation and Liability Act; CERCLA); *(5)* Organization of Economic Cooperation and Development (OECD, sublethal toxicity testing with fish and invertebrates); *(6)* European Economic Community (EC, sublethal toxicity testing with fish and invertebrates); and *(7)* the Paris Commission (behavioral endpoints).

5.1.4 Results of toxicity tests on sediments spiked at different concentrations of chemicals can be used to establish cause and effect relationships between chemicals and biological responses. Results of toxicity tests with test materials spiked into sediments at different concentrations may be reported in terms of an LC50 (median lethal concentration), an EC50 (median effect concentration), an IC50 (inhibition concentration), or as a NOEC (no observed effect concentration) or LOEC (lowest observed effect concentration). However, spiked sediment may not be representative of chemicals associated with sediment in the field. Mixing time Stemmer et al. 1990b, (**84**), aging (Landrum et al. 1989, (**85**), Word et al. 1987, (**86**), Landrum et al., 1992,(**87**)), and the chemical form of the material can affect responses of test organisms in spiked sediment tests.

5.1.5 Evaluating effect concentrations for chemicals in sediment requires knowledge of factors controlling their bioavailability. Similar concentrations of a chemical in units of mass of chemical per mass of sediment dry weight often exhibit a range in toxicity in different sediments Di Toro et al. 1990, (**88**) Di Toro et al. 1991,(**70**). Effect concentrations of chemicals in sediment have been correlated to interstitial water concentrations, and effect concentrations in interstitial water are often similar to effect concentrations in water-only exposures. The bioavailability of nonionic organic compounds in sediment is often inversely correlated with the organic carbon concentration. Whatever the route of exposure, these correlations of effect concentrations to interstitial water concentrations indicate that predicted or measured concentrations in interstitial water can be used to quantify the exposure concentration to an organism. Therefore, information on partitioning of chemicals between solid and liquid phases of sediment is useful for establishing effect concentrations Di Toro et al. 1991, (**70**).

TABLE 2 Sediment Quality Assessment Procedures (Modified from USEPA (78))

Method	Type			Approach
	Numeric	Descriptive	Combination	
Equilibrium Partitioning		*		A sediment quality value for a given contaminant is determined by calculating the sediment concentration of the contaminant that corresponds to an interstitial water concentration equivalent to the USEPA water-quality criterion for the contaminant.
Tissue Residues	*			Safe sediment concentrations of specific chemicals are established by determining the sediment chemical concentration that results in acceptable tissue residues.
Interstitial Water Toxicity	*	*	*	Toxicity of interstitial water is quantified and identification evaluation procedures are applied to identify and quantify chemical components responsible for sediment toxicity.
Benthic Community Structure		*		Environmental degradation is measured by evaluating alterations in benthic community structure.
Whole-sediment Toxicity And Sediment Spiking	*	*	*	Test organisms are exposed to sediments that may contain known or unknown quantities of potentially toxic chemicals. At the end of a specified time period, the response of the test organisms is examined in relation to a specified endpoint. Dose-response relationships can be established by exposing test organisms to sediments that have been spiked with known amounts of chemicals or mixtures of chemicals.
Sediment Quality Triad	*	*	*	Sediment chemical contamination, sediment toxicity, and benthic community structure are measured on the same sediment sample. Correspondence between sediment chemistry, toxicity, and field effects is used to determine sediment concentrations that discriminate conditions of minimal, uncertain, and major biological effects.
Sediment Quality Guidelines	*	*	*	The sediment concentration of contaminants associated with toxic responses measured in laboratory exposures or field assessments (that is, Apparent Effects Threshold (AET), Effect Range Median (ERM), Probable Effect Level (PEL).

5.1.6 Field surveys can be designed to provide either a qualitative reconnaissance of the distribution of sediment contamination or a quantitative statistical comparison of contamination among sites.

5.1.7 Surveys of sediment toxicity are usually part of more comprehensive analyses of biological, chemical, geological, and hydrographic data. Statistical correlations may be improved and sampling costs may be reduced if subsamples are taken simultaneously for sediment tests, chemical analyses, and benthic community structure.

5.1.8 Table 2 lists several approaches the USEPA has considered for the assessment of sediment quality USEPA, 1992, (**89**). These approaches include: *(1)* equilibrium partitioning, *(2)* tissue residues, *(3)* interstitial water toxicity, *(4)* whole-sediment toxicity and sediment-spiking tests, *(5)* benthic community structure, *(6)* effect ranges (for example, effect range median, ERM), and *(7)* sediment quality triad (see USEPA, 1989a, 1990a, 1990b and 1992b, (**90, 91, 92, 93** and Wenning and Ingersoll (2002 (**94**)) for a critique of these methods). The sediment assessment approaches listed in Table 2 can be classified as numeric (for example, equilibrium partitioning), descriptive (for example, whole-sediment toxicity tests), or a combination of numeric and descriptive approaches (for example, ERM, USEPA, 1992c, (**95**). Numeric methods can be used to derive chemical-specific sediment quality guidelines (SQGs). Descriptive methods such as toxicity tests with field-collected sediment cannot be used alone to develop numerical SQGs for individual chemicals. Although each approach can be used to make site-specific decisions, no one single approach can adequately address sediment quality. Overall, an integration of several methods using the weight of evidence is the most desirable approach for assessing the effects of contaminants associated with sediment, (Long et al. 1991(**96**) MacDonald et al. 1996 (**97**) Ingersoll et al. 1996 (**98**) Ingersoll et al. 1997 (**99**), Wenning and Ingersoll 2002 (**94**)). Hazard evaluations integrating data from laboratory exposures, chemical analyses, and benthic community assessments (the sediment quality triad) provide strong complementary evidence of the degree of pollution-induced degradation in aquatic communities (Burton, 1991 (**69**), Chapman 1992, 1997 (**100, 101**).)

5.2 *Regulatory Applications*—Test Method E1706 provides information on the regulatory applications of sediment toxicity tests.

5.3 *Performance-based Criteria:*

5.3.1 The USEPA Environmental Monitoring Management Council (EMMC) recommended the use of performance-based methods in developing standards, (Williams, 1993 (**102**). Performance-based methods were defined by EMMC as a monitoring approach which permits the use of appropriate methods that meet preestablished demonstrated performance standards (11.2).

5.3.2 The USEPA Office of Water, Office of Science and Technology, and Office of Research and Development held a workshop to provide an opportunity for experts in the field of sediment toxicology and staff from the USEPA Regional and Headquarters Program offices to discuss the development of standard freshwater, estuarine, and marine sediment testing procedures (USEPA, 1992a, 1994a (**89, 103**)). Workgroup participants arrived at a consensus on several culturing and testing methods. In developing guidance for culturing test organisms to be included in the USEPA methods manual for sediment tests, it was agreed that no one method should be required to culture organisms. However, the consensus at the workshop was that success of a test depends on the health of

the cultures. Therefore, having healthy test organisms of known quality and age for testing was determined to be the key consideration relative to culturing methods. A performance-based criteria approach was selected in USEPA, 2000 (**73**) as the preferred method through which individual laboratories could use unique culturing methods rather than requiring use of one culturing method.

5.3.3 This standard recommends the use of performance-based criteria to allow each laboratory to optimize culture methods and minimize effects of test organism health on the reliability and comparability of test results. See Annex A1 and Annex A2 for a listing of performance criteria for culturing or testing.

6. Interferences

6.1 *General Interferences:*

6.1.1 An interference is a characteristic of a sediment or a test system that can potentially affect test organism response aside from those related to sediment-associated contaminants. These interferences can potentially confound interpretation of test results in two ways: *(1)* toxicity is observed in the test sediment when contamination is low or there is more toxicity than expected, and *(2)* no toxicity is observed when contaminants are present at elevated concentrations or there is less toxicity than expected.

6.1.2 Because of the heterogeneity of natural sediments, extrapolation from laboratory studies to the field can sometimes be difficult (Table 3; Burton, 1991 (**69**)). Sediment collection, handling, and storage may alter bioavailability and concentration by changing the physical, chemical, or biological characteristics of the sediment. Maintaining the integrity of a field-collected sediment during removal, transport, mixing, storage, and testing is difficult and may complicate the interpretation of effects. See USEPA, 2000 (**64**) and Guide E1391. An abundance of the same organism (McGee et al., 1999 (**41**)) or organisms taxonomically similar to the test organism in the sediment sample may make interpretation of treatment effects difficult. In addition, the presence of predator may change the outcome of a toxicity test. For example, Redmond and Scott, 1989 (**104**) showed that the polychaete *Nephtys incisa* can consume *Ampelisca abdita* under toxicity test conditions. Similarly, predatory isopods (*Cyathura polita*) have been observed to interfere in 10-d toxicity tests conducted with *Leptocheirus plumulosus* (Peter De Lisle, Coastal Bioanalysts, Gloucester, VA; personal communication).

6.1.2.1 Although disruptive of natural sediment physical features, all test sediments in the *Leptocheirus plumulosus* 28-d sediment test should be press-sieved sometime before testing and re-homogenized immediately before introduction to the test chambers if warranted (section 10.3 and Annex A2). Press-sieving is performed primarily to remove predatory organisms, large debris, organisms used in testing (McGee et al., 1999 (**41**)) or organisms taxonomically similar to the test species. Certain applications may recommend that sediments should not be press-sieved. Also, it may not be necessary to press-sieve sediments if previous experience has demonstrated the absence of potential interferences, including predatory or competitive organisms or large debris, or if large debris or

TABLE 3 Advantages and Disadvantages for Use of Sediment Tests (Modified from Swartz (120))

Advantages
—Measure bioavailable fraction of contaminant(s).
—Provide a direct measure of benthic effects, assuming no field adaptation or amelioration of effects.
—Limited special equipment is required.
—Methods are rapid and inexpensive.
—Legal and scientific precedence exist for use; ASTM standards are available.
—Measure unique information relative to chemical analyses or benthic community analyses.
—Tests with spiked chemicals provide data on cause-effect relationships.
—Sediment-toxicity tests can be applied to all chemicals of concern.
—Tests applied to field samples reflect cumulative effects of contaminants and contaminant interactions.
—Toxicity tests are amenable to confirmation with natural benthos populations.
Disadvantages
—Sediment collection, handling, and storage may alter bioavailability.
—Spiked sediment may not be representative of field contaminated sediment.
—Natural geochemical characteristics of sediment may affect the response of test organisms.
—Indigenous animals may be present in field—collected sediments.
—Route of exposure may be uncertain and data generated in sediment toxicity tests may be difficult to interpret if factors controlling the bioavailability of contaminants in sediment are unknown.
—Tests applied to field samples may not discriminate effects of individual chemicals.
—Few comparisons have been made of methods or species.
—Only a few chronic methods for measuring sublethal effects have been developed or extensively evaluated.
—Laboratory tests have inherent limitations in predicting ecological effects.
—Tests do not directly address human health effects.

predators can be removed with forceps or other suitable tools. The presence of an abundance of amphipods that are taxonomically similar to the test species should prompt press-sieving. This is particularly true if endemic Ampeliscidae are present and *A. abdita* is the test species because it may be difficult to remove all of the resident amphipods from their tubes. If sediments are sieved, it is desirable to perform select analyses (for example, pore-water metals or DOC, AVS, TOC) on samples before and after sieving to document the influence of sieving on sediment chemistry (USEPA, 1994a (**1**)).

6.1.3 Depletion of aqueous and sediment-sorbed chemicals resulting from uptake by an organism or test chamber may also influence availability. In most cases, the organism is a minor sink for chemicals relative to the sediment. However, within the burrow of an organism, sediment desorption kinetics may limit uptake rates. Within minutes to hours, a major portion of the total chemical may be inaccessible to the organisms because of depletion of available residues. The desorption of a particular compound from sediment may range from easily reversible (labile; within minutes) to irreversible (non-labile; within days or months, Karickhoff and Morris, 1985 (**105**)). Interparticle diffusion or advection and the quality and quantity of sediment organic carbon can also affect sorption kinetics.

6.1.4 Testing sediments at temperatures different from the field might affect contaminant solubility, partitioning coefficients, or other physical and chemical characteristics.

Interaction between sediment and overlying water and the ratio of sediment to overlying water may influence bioavailability (Stemmer and Burton, 1990b (**84**)).

6.1.5 Results of sediment tests can be used to predict effects that may occur with aquatic organisms in the field as a result of exposure under comparable conditions. However, motile organisms might avoid exposure in the field. Photoinduced toxicity may be important for some compounds associated with sediment (for example, polycyclic aromatic hydrocarbons (PAHs) Davenport and Spacie, 1991 (**106**)). However, lighting typically used to conduct laboratory tests does not include the appropriate spectrum of ultraviolet radiation to photoactivate compounds (Oris and Giesy, 1985 (**107**), Ankley et al. 1994b (**108**)), and thus laboratory tests may not account for toxicity expressed by this mode of action.

6.1.6 Natural physico-chemical characteristics such as sediment texture may influence the response of test organisms (Dewitt et al. 1998, (**59**)). The physico-chemical characteristics of test sediment need to be within the tolerance limits of the test organism. Ideally, the limits of the test organism should be determined in advance; however, control samples reflecting differences in factors such as grain size and organic carbon can be evaluated if the limits are exceeded in the test sediment (section 12.1 and Annex A1 and Annex A2). The effects of sediment characteristics can also be addressed with regression equations Dewitt et al., 1998 (**59**) Ankley et al., 1994 (**60**). The use of formulated sediment can also be used to evaluate physico-chemical characteristics of sediment on test organisms (Walsh et al., 1991 (**61**), Suedel and Rodgers, 1994 (**62**)).

6.1.7 The route of exposure may be uncertain and data from sediment tests may be difficult to interpret if factors controlling the bioavailability of chemicals in sediment are unknown. Whole-sediment chemical concentrations may be normalized to factors other than dry weight. For example, concentrations of nonionic organic compounds might be normalized to sediment organic-carbon content, (USEPA, 1992 (**95**)) and certain metals normalized to acid volatile sulfides, (DiToro, 1990, (**88**)). Even with the appropriate normalizing factors, determination of toxic effects from ingestion of sediment or from dissolved chemicals in the interstitial water can still be difficult, (Lamberson and Swartz, 1998 (**109**)).

6.1.8 The addition of food, water, or solvents to the test chambers might obscure the bioavailability of chemicals in sediment or might provide a substrate for bacterial or fungal growth. Without addition of food, the test organisms may starve during long-term exposures (Ankley et al., 1994, McNulty et al. 1999 (**60, 110**)). However, the addition of the food may alter the availability of the chemicals in the sediment, (Harkey et al. 1994, Wiederholm et al. 1987 (**111,112**)) depending on the amount of food added, its composition (for example, total organic carbon (TOC)), and the chemical(s) of interest.

6.1.9 Laboratory sediment testing with field-collected sediments may be useful in estimating cumulative effects and interactions of multiple contaminants in a sample. Tests with field samples usually cannot discriminate between effects of individual chemicals. Many sediment samples contain a complex matrix of inorganic and organic chemicals with many unidentified compounds. The use of Toxicity Identification Evaluations (TIE) procedures including sediment tests with spiked chemicals may provide evidence of causal relationships and can be applied to many chemicals of concern (Ankley and Thomas, 1992, (**58**)). Laboratory studies that test single compounds spiked into the sediment can be used to determine more directly the specific chemicals causing a toxic response (Swartz et al. 1998 (**113**)).

6.1.10 Sediment spiking can also be used to investigate additive, antagonistic, or synergistic effects of specific chemical mixtures in a sediment sample (Swartz et al, 1998 (**113**)). However, spiked sediment may not be representative of contaminated sediment in the field. Mixing time (Stemmer et al. 1990a (**66**)), and aging (Landrum 1999, Word et al. 1997, Landrum and Faust 1992 (**85, 86, 87**) of spiked sediment can affect responses of organisms.

6.1.11 Salinity of the overlying water is an additional factor that can affect the bioavailability of metals. Importantly, some metals (for example, cadmium) are more bioavailable at lower salinities. Therefore, if a sediment sample from a low salinity location is tested with overlying waters of high salinity, there is the potential that metal toxicity may be reduced. The suite of species provided in this standard allow these tests to be conducted over the range of pore-water salinities routinely encountered in field-collected sediments from North American estuarine or marine environments (USEPA 1994a (**1**)). In addition, artificial sea salts may contain chelating agents (EDTA) that can potentially influence the bioavailability of metals. Certain brands of artificial salts are available from manufacturers without the addition of sodium thiosulfate that can also influence the toxicity of contaminants.

6.1.12 Most assessments of contaminated sediment rely on acute-lethality testing methods (for example, <10 d; (USEPA-USACE 1977, 1991, 1998, (**114, 115, 116**)). Acute-lethality tests are useful in identifying "hot spots" of sediment contamination, but may not be sensitive enough to evaluate moderately contaminated areas. Sediment quality assessments using sublethal responses of benthic organisms such as effects on growth and reproduction have been used to successfully evaluate moderately contaminated areas (Dillon et al. 1994, Kemble et al. 1994, Ingersoll and Brunson 1998, (**117, 82, 81**), Annex A2). Insufficient information is available to determine if the long-term test conducted with *Leptocheirus plumulosus* (Annex A2) is more sensitive than 10-d toxicity tests conducted with this or other species.

6.1.13 Despite the interferences previously listed, existing sediment testing methods that include measurement of sublethal endpoints may be used to provide a rapid and direct measure of effects of contaminants on benthic communities (for example, Canfield et al.. (**118**)). Laboratory tests with field-collected sediment can also be used to determine temporal, horizontal, or vertical distribution of contaminants in sediment. Most tests can be completed within two to four weeks. Legal and scientific precedence exist for use of sediment tests in regulatory decision making (for example, USEPA 1986a, Swartz 1989, (**119, 120**)). Furthermore, sediment tests with complex contaminant mixtures are important tools for making decisions about the extent of remedial action for

contaminated aquatic sites and for evaluating the success of remediation activities.

6.2 *Species-specific Interferences*—Interferences of tests for each species are described in Annex A1 and Annex A2.

7. Reagents and Materials

7.1 *Water:*

7.1.1 *Requirements*—Sea water used to test and culture organisms should be uniform in quality. Acceptable sea water should allow satisfactory survival, growth, or reproduction of the test organisms. Test organisms should not show signs of disease or apparent stress (for example, discoloration, unusual behavior). If problems are observed in the culturing or testing of organisms, it is desirable to evaluate the characteristics of the water. See USEPA (1993 (**121**)) and Guide E729 for a recommended list of chemical analyses of the water supply.

7.1.2 *Source:*

7.1.2.1 Culture and testing water can be natural or synthetic seawater (USEPA-USACE 2001 (**2**)).

7.1.2.2 The source of natural water will depend to some extent on the objective of the test and the test organism that is being used. All natural waters should be obtained from an uncontaminated surface-water source beyond the influence of known discharges. It may be desirable to collect water at slack high tide, or within one h after high tide. Suitable surface water sources should have intakes that are positioned to: *(1)* minimize fluctuations in quality and contamination, *(2)* maximize the concentration of dissolved oxygen (DO), and *(3)* ensure low concentrations of sulfide and iron. For estuarine tests, water having a salinity as near as possible to the desired test salinity should be collected from an uncontaminated area.

7.1.2.3 Alternatively, it may be desirable to dilute full strength sea water with an appropriate fresh water source. Sources of fresh water (that is, 0 ‰) for dilution include deionized water, uncontaminated well or spring water, or an uncontaminated surface-water source. Municipal-water supplies may be variable and may contain unacceptably high concentrations of materials such as copper, lead, zinc, fluoride, chlorine, or chloramines. Chlorinated water should not be used to dilute water utilized for culturing or testing because residual chlorine and chlorine- produced oxidants are toxic to many aquatic organisms. Dechlorinated water should only be used as a last resort for diluting sea water to the desired salinity since dechlorination is often incomplete (Guide E729; USEPA, 1993 (**121**)). It might be desirable or necessary to dilute full strength seawater with an appropriate freshwater source to achieve 5 % or 20 % (or the selected salinity; section 1.4) used in culturing or testing of *L. plumulosus* (USEPA-USACE 2001 (**2**), Section 12.

7.1.2.4 For site-specific investigations, it may be desirable to have the water-quality characteristics of the overlying water (that is, salinity) as similar as possible to the site water (section 1.4). For certain applications the experimental design might require use of water from the site where sediment is collected. In estuarine systems, however, the pore-water salinity of sediments may not be the same as the overlying water at the time of collection (Sanders et al., 1965 (**122**)).

7.1.2.5 Water that might be contaminated with facultative pathogens may be passed through a properly maintained ultraviolet sterilizer equipped with an intensity meter and flow controls or passed through a filter with a pore size of 0.45 µm or less.

7.1.2.6 Natural sea water might need aeration using air stones, surface aerators, or column aerators. Adequate aeration will stabilize pH, bring concentrations of DO and other gases into equilibrium with air, and minimize oxygen demand and concentrations of volatiles. The concentration of DO in source water should be between 90 to 100 % saturation to help ensure that DO concentrations are acceptable in test chambers. Natural sea water used for holding or acclimating, culturing, and testing amphipods should be filtered (<5 µm) shortly before use to remove suspended particles and organisms.

7.1.2.7 Water that is prepared from natural sea water should be stored in clean, covered containers at 4°C. USEPA-USACE (2001(**2**)) states that natural sea water should be used within 2 d for larval toxicity tests (Woelke, 1968 (**123**),1972 (**124**) ; Cardwell et al., 1977 (**125**),1979 (**126**)). However, investigators have found that when sea water is continuously aerated, it can be held for up to a month before use with certain species (David Moore, MEC Analytical, Carlsbad, CA; personal communication).

7.1.3 *Reconstituted/Synthetic Seawater:*

7.1.3.1 Although reconstituted water is acceptable, natural seawater is preferable, especially for tests involving chemicals whose bioavailability is affected by seawater chemistry. Reconstituted water can be prepared by adding specified amounts of reagent- grade chemicals to high-purity deionized water (Guide E729; USEPA, 1993 (**121**)). Acceptable high-purity water can be prepared using deionization or reverse-osmosis units (section 7.1; USEPA, 1993 (**121**)). Test water can also be prepared by diluting natural water with deionized water (Kemble et al., 1994 (**51**)).

7.1.3.2 Deionized water should be obtained from a system capable of producing at least 1 MΩ (mega-ohms) water. If large quantities of high quality deionized water are needed, it may be advisable to supply the laboratory grade water deionizer with preconditioned water from a mixed-bed water treatment system.

7.1.3.3 Reconstituted sea water is prepared by adding specified amounts of a suitable salt reagent to high-purity deionized water (Guide E729, USEPA, 1991(**127**)). Suitable salt reagents can be reagent grade chemicals, or commercial sea salts. Pre-formulated brine (for example, 60 to 90 %), prepared with dry ocean salts or heat-concentrated natural sea water, can also be used. (USEPA, 1994 (**1**) USEPA -USACE 2001, (**2**))

7.1.3.4 A synthetic sea formulation called GP2 is prepared with reagent grade chemicals that can be diluted with a suitable high-quality water to the desired salinity (USEPA, 1994b (**128**)).

7.1.3.5 The suitability and consistency of a particular salt formulation for use in holding and testing should be verified by laboratory tests because some formulations can produce unwanted toxic effects or sequester contaminants (Environment Canada, 1992 (**5**) ; USEPA-USACE 2001(**2**)). In controlled tests with the salt formulations mentioned above, Emery et al.