EPA United States Environmental Protection Agency

A Short-term Test Method for Assessing the Reproductive Toxicity of Endocrine-Disrupting Chemicals Using the Fathead Minnow (*Pimephales promelas*)

EPA/600/R-01/067
June 2002

A Short-term Test Method for Assessing the Reproductive Toxicity of Endocrine-Disrupting Chemicals Using the Fathead Minnow (*Pimephales promelas*)

U.S. Environmental Protection Agency

Office of Research and Development

National Health and Environmental Effects Research Laboratory

Mid-Continent Ecology Division

Duluth, MN 55804

DISCLAIMER

The information in this document has been funded wholly or in part by the U.S. Environmental Protection Agency. It has been subjected to review by the National Health and Environmental Effects Research Laboratory and approved for publication. Approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

FOREWORD

This manual describes a test method with the fathead minnow (*Pimephales promelas*) suitable for assessing potential reproductive effects of chemicals, with an emphasis on endocrine pathways controlled by estrogens and androgens. The test is conducted with reproductively-mature animals for 21 d. Endpoints assessed include: adult survival, reproductive behavior, secondary sex characteristics, gonadosomatic index, gonadal histology, plasma concentrations of vitellogenin and sex steroids (β-estradiol, testosterone, 11-ketotestosterone), fecundity, fertility, and, if desired, F_1 viability. In addition to describing the test method, guidance is presented as to interpretation of test results with respect to identification of potential endocrine-disrupting chemicals.

This document can be cited as EPA/600/R-01/067, U.S. Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Research Laboratory, Mid-Continent Ecology Division, Duluth, MN, USA.

CONTENTS

FIGURES

TABLES

ACKNOWLEDGMENTS

The basic test method described in this document is derived from research performed by scientists at the Mid-Continent Ecology Division (MED) of the National Health and Environmental Effects Research Laboratory (NHEERL) in the Office of Research and Development (ORD) of the U.S. Environmental Protection Agency (U.S. EPA). Portions of the document were prepared via a contract (8D-1337-NAEX) with the University of Wisconsin-Superior (Dr. Daniel Call, Mr. Larry Brooke). Histological guidance and related appendices were provided by Dr. Richard Leino, University of Minnesota-Duluth School of Medicine (IJ-0139-NTTX). Dr. Philip Dixon, Iowa State University, contributed to the statistical guidance (OJ-0248-NAEX). The lead authors of the document from MED are Dr. Gerald Ankley and Ms. Kathleen Jensen, with significant input by Dr. Michael Hornung, Mr. Michael Kahl, Mr. Joseph Korte, and Ms. Elizabeth Makynen. The authors gratefully acknowledge review comments from Dr. Tala Henry, Experimental Toxicology Division, NHEERL; and Dr. Thomas Hutchinson, Brixham Environmental Laboratory, AstraZeneca, Brixham, Devon, UK.

A. Scope

There has been recent concern for the potential effects of endocrine-disrupting chemicals (EDCs) on reproduction and development of humans and wildlife species (Colborn *et al.* 1996). The Office of Research and Development of the U.S. Environmental Protection Agency (U.S. EPA) has identified EDC issues as one of six high priority research areas (U.S. EPA 1996a; Kavlock *et al.* 1996; Ankley *et al.* 1997). Further, in response to legislation passed by the U.S. Congress (Food Quality Protection Act, PL 104-170; Safe Drinking Water Act, PL 104-182), the U.S. EPA is implementing a screening program for EDCs with specific mechanisms/modes of action (MOA). To aid in the development of this screening program, the U.S. EPA cosponsored a series of expert workshops on screening methods (Gray *et al.* 1997; Ankley *et al.* 1998a; DeVito *et al.* 1999), and convened a multi-stakeholder advisory committee (Endocrine Disruptor Screening and Testing Advisory Committee; EDSTAC) to recommend specific test methods and screening paradigms for EDCs (U.S. EPA 1998). The focus of these methods is on chemicals that may affect reproduction and/or development through disruption of physiological processes controlled by estrogen, androgen, and thyroid hormones. One Tier 1 screening assay recommended in U.S. EPA (1998) was a short-term (21 d) reproduction test with the fathead minnow (*Pimephales promelas*, Rafinesque) designed to identify chemicals that affect processes controlled by estrogens and androgens. A screening test with fish was considered particularly important for two reasons: (1) estrogenic/androgenic controls on reproduction/development in fish may differ significantly enough from that of higher vertebrates such that mammalian (rat) screening methods may not identify potential EDCs in this important class of animals, and (2) as

opposed to human health effects, there is currently significant evidence of adverse EDC effects in a variety of wildlife species, including fish (Crisp *et al.* 1997; Ankley and Giesy 1998). The purpose of this document is to describe an EDC test method with the fathead minnow designed to meet the requirements of U.S. EPA (1998).

B. Introduction

The fathead minnow was selected as the test organism for this reproduction assay for a number of reasons. Attractive attributes of this species include its: (1) widespread geographical distribution, (2) representation of an ecologically-important family of fish (Cyprinidae), (3) rapid development and sexual maturation, (4) ease of culturing and testing, (5) common use as a warm-water species in regulatory testing and decision-making, and (6) the existence of extensive chemical toxicological databases including information on reproductive physiology and endocrinology that is valuable in the context of monitoring EDCs.

1. Basic Reproductive Biology

The fathead minnow is an omnivorous freshwater fish in the family Cyprinidae. It is a relatively hardy species with a broad geographic distribution across North America (Devine 1968; Held and Peterka 1974; U.S. EPA 1987). The fathead minnow has a relatively rapid life-cycle, achieving reproductive maturity within four to five months of hatch (under optimal conditions). The timing of the reproductive cycle can be controlled effectively through the

use of temperature and photoperiod manipulation (U.S. EPA 1987), thus enabling a lab to maintain a constant supply of test organisms at a developmental stage suitable for testing. At maturity, males weigh 4 to 5 g, and females weigh 2 to 3 g. Spawning can successfully occur at pH values ranging from 6.6 to 9.5 (Mount 1973). Fathead minnows tolerate total alkalinity concentrations of up to 1,800 mg/L as $CaCO_3$ (McCarraher and Thomas 1968), and turbidity as high as 15,000 mg/L total solids (Rawson and Moore 1944). The species also is tolerant of water temperatures ranging from 2 to 33°C (Bardach *et al.* 1966), and spawns successfully in the temperature range of 15.6 to 29.8°C (Brungs 1971).

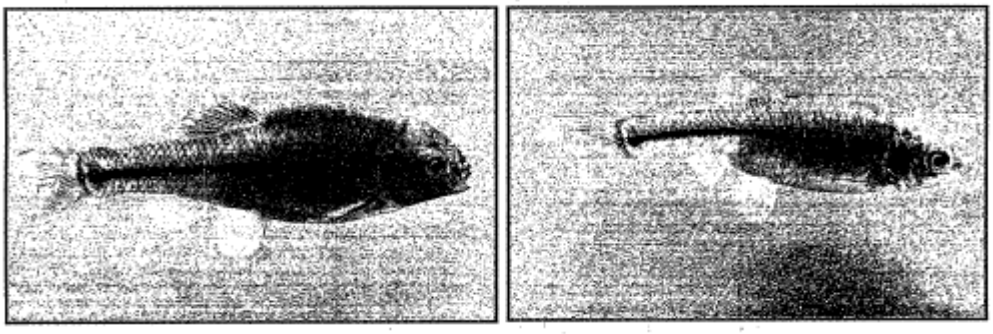

Figure 1. Mature male (left) and female (right) fathead minnow, *Pimephales promelas*.

The adult fathead minnow is sexually dimorphic, with males and females readily distinguishable from one another when in breeding condition (Fig. 1). As juveniles, the sexes are similar in appearance. Sexually-mature males develop large nuptial tubercles on the

snout, and an elongated, fleshy, dorsal pad which extends as a narrow band from the nape to the dorsal fin. The body coloration of males becomes black on the sides except for two wide light-colored vertical bars (U.S. EPA 1987). In contrast, females generally do not undergo obvious changes in color or morphology, but do develop a fleshy ovipositor that can be used to definitively distinguish females from immature/quiescent males (Flickinger 1969). Studies have demonstrated that secondary sex characteristics in this species are under the control of sex steroids and, hence, could be affected by chemicals such as estrogen or androgen receptor agonists or antagonists (Smith 1974; Miles-Richardson *et al.* 1999a; Harries *et al.* 2000; Ankley *et al.* 2001; Länge *et al.* 2001). It should be noted, however, that there can be some degree of ambiguity in differentiating sex even in ostensibly mature fathead minnows. For example, less dominant males from mass culture situations may resemble females phenotypically. Alternatively, we also have noted a baseline incidence of reproductively-active females whose barred coloration patterns can resemble that observed in males. This seems to occur in situations (such as in the test described herein) where there are multiple females per test tank; interestingly, in these situations a hierarchal behavior can develop in which there is a dominant female in the tank which can exhibit male coloration patterns (MED, unpublished data).

Breeding males are territorial and seek out nest sites which they actively defend against other males and intruders. Fathead minnows spawn beneath objects (artificial substrates in the lab) and the buoyant, adhesive eggs stick to each other and to the undersurface of the nesting substrate. Spawning behavior is characterized by close lateral contact and body vibration

between male and female. After sufficient stimulation, the male presses the female upward, resulting in the female's urogenital region contacting the substrate, with a concomitant release of eggs. Milt is released at this time as the pair terminates pressing with an abrupt separation. This behavior occurs intermittently until an increase in male aggression drives the female away. The male then guards and tends the nest. The latter activity includes cleaning the eggs of detritus and agitating the water around the eggs, thereby ventilating them with oxygenated water. At 25°C, the embryos hatch in about 4 to 5 d (U.S. EPA 1996b). The number of eggs per spawn can be variable, depending upon the age, size, and condition of the female (Gale and Buynak 1982). However, the mean number of eggs per spawn under stable laboratory conditions is typically in the range of 50 to 150 (Jensen *et al.* 2001). Under the test conditions described herein (25 ± 1°C, 16:8-h light:dark photoperiod), individual females spawn at intervals of 3 to 4 d (Jensen *et al.* 2001), creating the potential for a single female to routinely produce in excess of 500 eggs during a 21-d test.

2. Toxicology

The fathead minnow has been used extensively in short-term (acute) and long-term (partial life-cycle or complete life-cycle) chemical toxicity studies. They have been widely tested as a representative warm-water species to provide acute and chronic toxicity data for the preparation of U.S. EPA national ambient water quality criteria documents. In addition, fathead minnow toxicity tests with more than 600 chemicals form a unique archival database (Brooke *et al.* 1984; Geiger *et al.* 1985; 1986; 1988; 1990; Mayer and Ellersieck 1986; Call

and Geiger 1992) that has been used extensively for a number of purposes, including quantitative structure-activity relationship (QSAR) modeling (Russom *et al.* 1991; 1997). Although the fathead minnow has been tested occasionally in full life-cycle assays incorporating a variety of reproductive and developmental endpoints (McKim 1977; Länge *et al.* 2001), shorter-term tests have been more typical for this species. For example, protocols are available for 4-d survival and 7-d survival and growth tests that start with either newly hatched larvae or embryos (U.S. EPA 1993; 1994). Protocols for slightly longer early life-stage tests of ≥28 d are also available; these typically are initiated with embryos, and include exposure of the embryos, newly hatched fry, and/or juvenile fish to some point prior to sexual maturation (ASTM 2000b). A comparison of short- versus long-term toxicity tests with the fathead minnow (McKim 1977), illustrated the utility of early life-stage tests of 1 to 4 months in duration in providing an estimate of chemical toxicity over a complete life-cycle. However, for the purposes of ecological risk assessment, it has been recommended that some measure of reproduction be incorporated into the partial life-cycle tests (e.g., Suter *et al.* 1987). None of the above-mentioned short-term protocols address possible effects on reproduction in adult fish. Therefore, the test protocol described herein includes several direct and indirect measures of fecundity as endpoints. This basic test protocol also can capture endpoints assessed by traditional partial life-cycle and early life-stage protocols (Ankley *et al.* 2001).

From an ecological perspective, determination of effects of toxicants on reproductive fitness and, hence, possible population-level impacts clearly is critical (Suter *et al.* 1987). However,

in terms of screening for chemicals that cause toxicity via MOA of known concern, endpoints specific to these pathways also are important. In recognition of this, endpoints suggested as desirable for EDC screening in fish models include effects on reproductive behavior, secondary sex characteristics, gonadosomatic index, gonadal histology, and plasma concentrations of vitellogenin and sex steroids (β-estradiol, testosterone, 11-ketotestosterone) (Ankley *et al.* 1998a; 2001; U.S. EPA 1998; Jensen *et al.* 2001). Recent studies have assessed the use of these endpoints in EDC studies with the fathead minnow, confirming their utility in this species, and providing important baseline data in terms of interpretation of results obtained from these standard test protocols. Specifically, induction of vitellogenin in response to estrogen receptor agonists (Kramer *et al.* 1998; Panter *et al.* 1998; 2002; Parks *et al.* 1999; Tyler *et al.* 1999; Harries *et al.* 2000; Korte *et al.* 2000; Ankley *et al.* 2001; Länge *et al.* 2001), and alterations in gonadal histology or secondary sex characteristics associated with exposure to estrogen or androgen receptor agonists (Smith 1974; Miles-Richardson *et al.* 1999a,b; Harries *et al.* 2000; Ankley *et al.* 2001; Länge *et al.* 2001), have been characterized in EDC screening studies with the fathead minnow. Finally, there also is an emerging database concerning the effects of EDCs with known MOA on patterns of circulating sex steroids in this species (Giesy *et al.* 2000; Makynen *et al.* 2000; Ankley *et al.* 2001; 2002; Jensen *et al.* 2002).

C. Acronyms and Definitions

Acute toxicity - Effects observed in tests of ≤96 h in duration.

AQUIRE - AQUatic Information REtrieval system; a U.S. EPA database of aquatic toxicity information.

ASTER - ASsessment Tools for the Evaluation of Risk; a U.S. EPA software program that combines aquatic toxicity databases and quantitative structure-activity relationship toxicity prediction models to assist in the development of risk assessments for chemicals.

Chronic toxicity - Effects observed in fathead minnow tests ≥28 d in duration.

Dorsal pad - Soft enlargement of flesh on top of the head of sexually-mature male fathead minnows that extends onto the back of the fish to, or near, the anterior margin of the dorsal fin.

EDC - Endocrine-disrupting chemical.

ELISA - Enzyme-Linked Immunosorbent Assay; analytical method used for determining plasma vitellogenin concentration.

β-Estradiol - Major estrogenic sex steroid regulating reproductive function.

Fecundity - Measure of total egg production.

Fertility - Measure(s) of fertilization success as indicated, for example, by actively-dividing embryonic cells or occurrence of eyed embryos.

GSI - Gonadosomatic Index; gonad weight relative to total body weight ((gonad wt(g)/body wt (g)) x 100).

i.p. injection - Intraperitoneal injection; method of chemical delivery.

11-Ketotestosterone - Major male sex steroid in fish responsible for development of secondary sex characteristics as well as gonadal development.

LC50 - Concentration lethal to 50% of a group of organisms under specified conditions.

LOEC (LOAEC) - Lowest Observed Effect Concentration; lowest concentration of a chemical that causes a significant adverse effect upon one or more life functions.

MOA - Mechanism/Mode of Action; the mechanism or mode via which a chemical exerts a toxic response in an organism.

NOEC (NOAEC) - No Observed Effect Concentration; highest concentration of a chemical that does not cause a significant adverse effect upon any life functions.

Nuptial tubercles - Visible external horny outgrowths on the surface of the head of the sexually-mature male fathead minnow in breeding condition.

Ovipositor - Urogenital structure present in sexually-mature females for egg deposition.

QSAR - Quantitative Structure-Activity Relationship; a relationship between basic chemical structure (or property) and a biological response that is described quantitatively.

RIA - RadioImmunoAssay; analytical method used for determining plasma steroid concentrations.

RPD - Relative Percent Difference; calculation utilized to assess measurement precision.

Saturator - An apparatus capable of generating a saturated stock solution of a chemical that is relatively insoluble in water.

Subchronic toxicity - Effects observed in fathead minnow tests of >4- and <28 d in duration.

Testosterone - Androgenic sex steroid normally present in both sexes and necessary for development and maintenance of reproductive function.

Viability - Measure(s) of embryonic development subsequent to fertilization, including hatching success and normal larval maturation.

Vitellogenin - Precurser to egg yolk protein that occurs normally in the blood of sexually-mature female fish; it can be induced by estrogen receptor agonists in male fathead minnows.

D. Principle of the Test

This test is designed as a short-term reproduction assay suitable for identifying chemicals that affect reproduction or, potentially, development through disruption of any of a number of pathways, including those controlled by estrogens and/or androgens. Several potentially sensitive endpoints are assessed. The test is initiated with mature male and female fish that have a documented history of reproductive success as measured both by fecundity (number of eggs) and by embryo viability (e.g., hatch). This is established during a pre-exposure phase of 14 to 21 d in the same system/test chambers as will be utilized for the chemical exposure. During the subsequent 21-d chemical exposure, survival, reproductive behavior, and secondary sex characteristics are observed, and fecundity (number of spawns and number of eggs/spawn) monitored daily. Viability of resultant embryos (e.g., hatching success, developmental rate, occurrence of malformations) can be assessed in animals held either in clean water or in the same treatment regime to which the adults were exposed. At conclusion of the 21-d test, blood samples are collected from the adults for determination of plasma vitellogenin and sex steroids, and the gonads sampled for measurement of the gonadosomatic index (GSI) and histological analyses. Effects in the treatment groups are assessed by comparison to control groups to determine if any of the endpoints in the exposed fish are significantly different from those in

controls. Those chemicals identified as positive in this test may be advanced for more comprehensive testing.

E. Information on the Test Chemical

1. Toxicity to Fish

At least two concentrations of the test chemical should be evaluated in the 21-d reproduction assay, but additional concentrations are desirable if resources are available (SETAC 1997). It is also required that the test be conducted as a flow-through, rather than a static-renewal assay, and that concentrations of the chemical of concern be analytically determined in the stock solution and test chambers during testing. In addition, it would be desirable to measure concentrations of the test chemical and relevant metabolites in tissue of the fish at conclusion of the test. The highest exposure concentration used in the test should be one that is not lethal to the fish and can be maintained at a near constant level over the duration of the exposure period. When using a waterborne delivery route this concentration generally should not exceed water solubility. Depending upon data available for the proposed test chemical, the following can be used to determine the highest exposure concentration. The order of priority in terms of utility of available information for establishing this concentration is: fathead minnow complete life-cycle toxicity data > fathead minnow early life-stage toxicity data > adult fathead minnow acute toxicity data > chronic toxicity data with surrogate fish species. Performance of a range-finding toxicity test of 4 to 7 d in length with adult fathead

minnows is strongly recommended, particularly if toxicity data are lacking for the fathead minnow. Fish toxicity data estimated from QSAR models may be used to determine concentrations to use in a range-finding toxicity test (described below) when no empirical test data are available for the compound of interest.

To initially ascertain if toxicity data exist for a given test substance a literature search should be conducted. Possible sources of toxicity information include chemical-specific U.S. EPA National Ambient Water Quality Criteria or Advisory documents (U.S. EPA, Office of Water, Washington, DC). These documents describe tests that have generated data of established quality; they should be consulted to determine if chronic toxicity data for the chemical(s) of concern exist for the fathead minnow. Another source of acute and chronic toxicity data is the AQUIRE (AQUatic Information REtrieval system) database. This database is accessible as a component of the larger ECOTOX database (web address: www.epa.gov/ecotox/). Additional information on accessing this database is available by e-mail at the following address: ecotox.support@epa.gov.

Data of the highest quality should be used in selecting test concentrations. For example, if data are available for both flow-through and static tests, the flow-through data generally should be utilized. Similarly, if data are available for assays in which chemical concentrations were measured versus tests in which they were not, the former usually should be used.

If chronic data are available for the fathead minnow, the highest concentration used should be the LOEC (lowest observable effect concentration), if the adverse effect was upon growth or reproduction; or the NOEC (no observable effect concentration), if the LOEC was based upon survival.

If acute, but not chronic, toxicity data for the test chemical are available for the fathead minnow a divisor may be used to estimate a chronic value from the 96-h LC50 value. A divisor of approximately 10 to 12 has been found to be representative of the mean difference between the 96-h LC50 value and the chronic value (geometric mean of NOEC and LOEC) for chemicals which have a narcosis MOA (Kenaga 1982; Call *et al.* 1985). For chemicals with a more specific MOA, a larger divisor may be appropriate. Selection of a maximum test concentration from acute toxicity information ideally should be augmented by a range-finding test prior to the 21-d reproduction test.

If acute or chronic toxicity data are not available for the proposed test chemical and the fathead minnow, information should be sought for other fish species. Again, one source of toxicity data that may prove useful is the AQUIRE database. Such information could be used to establish a maximum concentration for a range-finding test with the fathead minnow. If only acute toxicity data are available, an appropriate divisor should be applied as described above. In either case, a range-finding test with the fathead minnow should precede the 21-d reproduction assay.

If there are no toxicity data available for any fish species, but data exist for chemicals from the same class which are known or presumed to act by the same toxic MOA as the proposed test chemical, acute toxicity may be predicted using a QSAR model. Models for various classes of chemicals are available, as well as some general rules for model selection (Bradbury and Lipnick 1990; Veith and Broderius 1990; Bradbury 1994; 1995; Russom *et al.* 1997). Based on this information, a range-finding test should be conducted to help identify appropriate test concentrations for the 21-d reproduction assay.

Lower test concentrations of the chemical of concern will generally be comprised of a fractional geometric progression relative to the highest concentration used in the assay; the magnitude of this fraction will depend somewhat on the number of concentrations utilized. For example, if only two or three concentrations are used in the test, the lower concentration(s) should be a factor of 5 to 10 times less than the highest concentration. When utilizing a larger number of test chemical concentrations (e.g., ≥5), a 50% dilution series is relatively common, although a logarithmic concentration series is not unusual (e.g., 100, 32, 10, 3.2, etc.). As a general rule, when sensitivity of the analytical chemistry method for the material of concern is an issue, the lowest concentration tested should be near the detection limit and higher concentrations should range uniformly to the highest test concentration.

2. Physico-Chemical Properties and Chemical Delivery

It is necessary to obtain reliable information concerning the physico-chemical properties of test chemicals to determine likely behavior in the test system and appropriate exposure methodology. Factors important in this regard might include water solubility, octanol-water partition coefficient (K_{ow}), melting point, density, volatility, stability in water, and biodegradability. Information on these characteristics may be obtained from single sources such as Lewis (1991) or Keith (1997), literature searches, or computerized programs, such as ASTER (ASsessment Tools for the Evaluation of Risk; Russom *et al.* 1991). ASTER is a UNIX-based computer program that is not publicly available. To obtain chemical property information from ASTER, contact should be made via initial e-mail at: ecotox.support@epa.gov. Other chemical property databases are commercially available. For example, the database ENVIROFATE is available through National Technical Information Service, Springfield, VA (www.ntis.gov); and the databases ECDIN (Environmental Chemicals Data & Information Network), EFD (Environmental Fate Database), LOGKOW, QSAR, and MicroQSAR are available through Technical Database Services, Inc., New York, NY (www.tds-tds.com).

Knowledge of these physico-chemical properties will allow determination of an appropriate procedure to use in producing a stock solution of the test chemical, as well as the necessary rate of renewal (i.e., derived from aqueous stability). For example, if the chemical is highly soluble in water (≥1,000 mg/L), stock solutions may be prepared directly in the test water

with a slow-stirring method. High-energy stirring, such as with a mechanical blender, generally should be avoided, as it may produce emulsified droplets of pure compound that could enter the exposure system. The stock solution may then be pumped directly into a holding chamber in the test system for subsequent dilution. If the chemical is relatively insoluble in water, and a liquid at the target test temperature (i.e., melting points <25°C), a liquid-liquid saturator is recommended for generating stock solutions; whereas, if the chemical is a solid, a glass wool column saturator is the preferred method of stock solution generation (Kahl *et al.* 1999). These methods are described in greater detail in Section G.6.

None of the above methods require solvent or oil carriers. This is preferred because: (1) certain solvents themselves may result in toxicity and/or undesirable or unexpected endocrinological responses, (2) testing chemicals above their water solubility (as can frequently occur through the use of solvents) can result in inaccurate assessments of risk from the perspective of contaminant bioavailability, and (3) the use of solvents in longer-term tests can result in a significant degree of "biofilming" associated with microbial activity. If, however, it is determined that a solvent carrier must be used, several choices are possible including acetone, dimethylsulfoxide, dimethylformamide, ethanol, ethylene glycol, and methanol. Toxicity and/or potential endocrinological effects of these solvents have not been established for the fathead minnow in subchronic or chronic assays; however, all are of relatively low toxicity in 96-h lethality tests (Brooke *et al.* 1984; Phipps and Holcombe 1985; Poirier *et al.* 1986; Geiger *et al.* 1990; Pillard 1995). If solvent carriers are used, appropriate solvent controls must be evaluated in addition to non-solvent controls. If it is not possible to

administer a chemical via the water, either because of physico-chemical characteristics (low solubility) or limited chemical availability, it may be necessary to introduce the chemical via the diet or by intraperitoneal (i.p.) injection (e.g., Korte *et al.* 2000; Kahl *et al.* 2001). These routes are less preferable than aqueous administration; this is, in part, due to the fact that dietary and i.p. routes have not been commonly used with the fathead minnow, which results in uncertainty in the current identification as to optimized protocols.

3. Range-Finding Toxicity Tests

Given the fact that there will be little acute or chronic toxicity data for many chemicals, or that existing data are unreliable (e.g., generated from static tests with unmeasured concentrations), range-finding toxicity tests often will be needed to define appropriate exposure concentrations for the 21-d reproduction test. The highest concentration for a range-finding test could be derived from toxicity data for other fish species and/or a QSAR model (see above). In the absence of any type of empirical or predicted toxicology data, the range-finding test should start at the solubility limit of the chemical in water. Test concentrations should decrease by a factor of 10 for each successively lower exposure level. The range-finding test should ideally be performed with reproductively-mature fathead minnows under conditions (e.g., fish age, loading rate, temperature, chemical source/delivery) similar to those to be used in the 21-d reproduction test (described in detail in Sections G and H, and briefly summarized in Table 1). Ideally, a minimum of two test chambers (replicates) per concentration, each containing four females and two males, should

be exposed in the range-finding test. Alternative designs of range-finding tests with less replication could possibly be considered on a case-by-case basis. The exposure period should be a minimum of 4 to 7 d, with a longer period preferable. The number of mortalities that occur, and the nature of the concentration (dose)-response curves over time of exposure can provide critical information in determining the maximum concentration to use in the 21-d reproduction test. Based on these results, the highest range-finding test concentration that does not result in increased mortality or signs of overt morbidity (e.g., cessation of feeding), compared to the controls should serve as the highest exposure concentration in the 21-d test.

4. Analytical Determinations

A literature review should be conducted to identify analytical methods that have been used to measure concentrations of the test substance in water (or, depending on nature of the exposure route, in food or tissue). An appropriate method should be selected and verified prior to performance of the range-finding and/or 21-d reproduction test.

F. Validity of the Test

The 21-d reproduction test can be considered valid only if certain conditions are met. These include: (1) documentation of health of the test animals, as determined from their survival and reproductive performance during a pre-exposure period, (2) high survival (≥90%) of parental control fish during the exposure period, (3) active spawning (egg production) of parental control

fish, (4) a high rate (≥90%) of fertility in the control group, (5) maintenance of water quality characteristics (i.e., temperature, dissolved oxygen, pH, alkalinity, total organic carbon, and un-ionized ammonia) within specified limits (Table 2), (6) successful analytical measurement of the test chemical concentrations in the exposure media (e.g., water for aqueous exposures), and (7) maintenance of chemical exposure concentrations within specified limits (see Section H.2.e for limits).

G. Description of the Method

An overview of the test conditions is provided in Table 1. Specific aspects of the exposure system(s)/test conduct are presented below.

1. Exposure Apparatus

There is no absolute requirement for a particular physical apparatus for the test described herein. Water needs to be delivered to the fish in a consistent manner with controlled concentrations of the chemical of concern, but several options are available. For example, the apparatus for flow-through tests may be modified from a proportional diluter system, of which several designs are available (Mount and Brungs 1967; Defoe 1975; Benoit *et al.* 1982). Water enters the top of the unit and proceeds through the proportional dilution system and exposure chambers via gravity feed. Other systems may operate via pumps, as opposed to gravity feed and be controlled, for example, by electronic units (e.g., DeFoe and Holcombe

1997). Water initially enters the system from a common source (same source as the water used to culture and/or precondition the fish prior to chemical exposure) and is distributed in equal portions to each replicate control and treatment cell. The water supply must be adequate to provide a minimum of 960 L (10 L/4 h x four replicates of control, solvent control, and two chemical treatments) of water daily to the test system (approximately 3.5 L/g fish/day). Some testing may require a greater capacity for water delivery (e.g., three or more chemical concentrations). Further details concerning test systems/chemical delivery are given below.

2. Test Chambers

Glass, stainless steel, or other chemically-inert materials should be used for exposure of the fish to test solutions. Materials with the potential to leach potential endocrine-active substances such as phthalate esters should be avoided. The dimensions of the test chambers must be such that reproduction occurs at a consistent rate comparable, for example, to that achieved under culture conditions. To ensure active spawning and successful fertilization, four females and two males per replicate exposure chamber are currently recommended. The minimum recommended chamber size is 40 cm long, 20 cm wide, and 20 cm high containing 10 L of test solution. Each test chamber should contain three spawning substrates constructed of stainless steel or PVC pipe (10-20 cm in length) split lengthwise (U.S. EPA 1987). Inflow of water to the chamber should be at the end opposite the outflow to help ensure a complete replacement with test/clean water.

Table 1. Overview of recommended exposure conditions for the fathead minnow 21-d reproduction test.

1. Test type	Flow-through
2. Water temperature	25 ± 1°C
3. Illumination quality	Fluorescent bulbs (wide spectrum)
4. Light intensity	10-20 $\mu E/M^2/s$, 540-1080 lux, or 50-100 ft-c (ambient laboratory levels)
5. Photoperiod	16 h light, 8 h dark
6. Test chamber size	18 L (40 x 20 x 20 cm) (minimum)
7. Test solution volume	10 L
8. Volume exchanges of test solutions	Minimum of six daily
9. Flow rate	Approximately 3.5 L/g fish/day
10. Age of test organisms	Reproducing adults (120 d minimum)
11. No. of fish per test chamber	Four females and two males
12. No. of treatments	Two minimum (plus appropriate controls)
13. No. of replicates per treatment	Four minimum
14. No. of fish per test concentration	Minimum of 16 females and 8 males
15. Feeding regime	Frozen adult brine shrimp twice daily if toxicant exposure is aqueous or via i.p. injection
16. Aeration	None unless dissolved oxygen concentration falls below 4.9 mg/L
17. Dilution water	Clean surface, well, or reconstituted water
18. Chemical dilution factor	5-10
19. Chemical exposure duration	≤ 21 d
20. Primary endpoints	Adult survival, reproductive behavior, secondary sex characteristics, gonadosomatic index (GSI) and gonadal histology, plasma vitellogenin and sex steroid (β-estradiol, testosterone, 11-ketotestosterone) concentrations, fecundity, and fertility
21. Optional endpoints	Embryo hatch, larval survival, and morphology
21. Test acceptability	Dissolved oxygen ≥60% of saturation; mean temperature of 25 ± 1°C; 90% survival in the controls; successful egg production in controls

3. Selection of Test Species

The fathead minnow is used for this test. The rationale for selecting this species has been described previously which, in summary, is due to their ease of culturing, short life-cycle, and history of use in toxicity testing.

4. Test Water

Any surface, well, or reconstituted water that results in acceptable survival and reproductive viability is suitable. The water supply should be adequate to maintain constant conditions of water quality during both culture and testing (Table 2). The characteristics of the water ideally should not impact the availability of the test chemical to the test organisms; thus, waters high in dissolved or particulate organic carbon should be avoided. Periodic (e.g., yearly) sampling and analysis should be made of the dilution water for potentially toxic metals (e.g., Cu^{++}, Ni^{++}, Zn^{++}, Pb^{++}, Hg^{++}, Cd^{++}), major cations (e.g., Ca^{++}, Mg^{++}, Na^{+}, K^{+}), major anions (e.g., Cl^{-}, S^{--}), priority pesticides, total organic carbon, suspended solids, and nitrates.

Table 2. Recommended ranges of water quality characteristics for fathead minnow reproduction.

Water Characteristic	Recommended Range	Reference
Temperature (°C)	24.0-26.0	ASTM (2000a,b)
Dissolved oxygen (mg/L)	>4.9 mg/L (≥60% saturation)	ASTM (2000a,b)
pH	6.5-9.0	U.S. EPA (1976)
Alkalinity (mg/L as $CaCO_3$)	>20	U.S. EPA (1976)
Total organic carbon (mg/L)	≤5	ASTM (2000b)
Un-ionized ammonia (μg/L)	<35	ASTM (2000a)

5. Test Type

The animals must be tested using a flow-through system. The water must be used only once before discharge to waste. Some test chemicals may require that exposure water be treated prior to discharge to municipal wastewater treatment plants. For the purposes both of consistent (parent) chemical exposure and maintenance of adequate water quality (e.g., dissolved oxygen, ammonia, etc.), it is desirable to have the test solution reside in the chamber with the animals for a limited amount of time. The test solution must have a renewal rate of at least one volume exchange every 4 h, but certain chemicals will need to be

renewed more frequently due to their volatility, degradation, or complexation with the test containers.

6. Stock Solutions (Aqueous Exposures)

Chemical stock solutions can be generated using a number of approaches, and delivered via either proportional dilution systems or stock pumping systems (Mount and Brungs 1967; DeFoe 1975; Benoit *et al.* 1982; DeFoe and Holcombe 1997). The method of choice for preparation of stock solution(s) must provide for an adequate volume of water to achieve desired test concentrations for a flow-through test. It is preferable that stock solutions be prepared in water without the use of solvents or solubility enhancers. As described below, different approaches have been used to successfully achieve appropriate test concentrations for chemicals of various physico-chemical characteristics in the absence of solvents. In some instances, the use of a solvent for chemical delivery in waterborne exposures will be unavoidable.

a. Solid-Liquid (slow-stir) Saturator. A solid-liquid saturator is designed to use the surface of the inside of a glass carboy or vessel as an area which, when coated with a test chemical, generates a stock solution at or near saturation. This is a slight adaptation of a system in which stock solutions are prepared directly by dissolving test chemical in water. Compounds which are solids and moderately soluble in water are best suited for this method. Generally, these types of chemicals are not soluble

enough to mix in stock vessels for immediate use, but not so insoluble that it is necessary to coat substrates with a large surface area, such as a glass wool column saturation unit described below. To determine the effective use of this method in a test system, analytical measurements should be taken from the chemically-coated vessel at a range of flow rates over time.

The test chemical is dissolved in a solvent (e.g., acetone, ethanol) before application on the glass vessel ($\geq$ 1L), with the amount of solvent varying due to vessel size and amount of chemical. Coating of the glass with the chemical requires an even distribution, which can be successfully achieved by rolling the container under a low flow of air or nitrogen gas to evaporate solvent from the system. Vessels that can be placed on a rolling mill simplify coating the chemical evenly. Upon solvent dryness, usually identified by a crystalline design, the vessel is filled with test water to the desired level, fitted with the stopper assembly and secured. It is important that the air space in the saturator be completely sealed from air exit or entry during use of the saturator, and that it be possible to monitor the rate of incoming make-up water at the rate of the diluent being removed. Water input occurs just above the water level in the vessel, with an output tube located near the bottom. The diameter of the output tubing is best when matched with that used on the dispensing pump. Rapid stirring within the vessel is produced using a magnetic stir bar and stir motor. The presence of a vortex at the air/liquid interface generally is sufficient for adequate stock mixing. This type of saturator has been used successfully with the model antiandrogen

flutamide (Jensen *et al.* 2002). In that study, a 3.5 g quantity of chemical was dissolved in 40 ml of acetone. With an 8 ml/min flow rate through an 18 L carboy, stock concentrations were generated at about 21 mg/L for up to 7 d. Therefore, three coated vessels supplied a usable stock for the 21 d of chemical exposure.

b. Liquid-Liquid (slow-stir) Saturator. In a liquid-liquid system, water that is nearly saturated with the test compound is pumped from the saturator vessel to the diluter system, and at the same time replacement water is added back to the saturator at the rate that it was withdrawn. This type of saturator has been successfully used with chemicals that are liquid at room temperature with water solubilities between 3.2 and 20,000 mg/L (Kahl *et al.* 1999). A glass container of sufficient volume is required to achieve a surface to volume ratio between chemical and water that is adequate to generate a volume of test solution near water saturation for the chemical. The system generally consists of a glass container (≥ 1 L) containing water, test chemical, a magnetic stir bar, and air space (Fig. 2). The container should have a narrow neck into which a neoprene stopper can fit. Four stainless steel or glass tubes are inserted through the stopper to permit input of chemical, venting, input of dilution water, and output of the test solution. If the test chemical has a specific gravity <1, it will float on the surface of the water and the dilution water input tube should be positioned above the solution level, and the output tube near the bottom of the container (Fig. 2A). If the specific gravity of the test chemical is >1, the chemical will be close to the bottom of the container. In this case, the dilution water input tube should be

near the bottom, and the output tube in the upper half of the solution (Fig. 2B). The container is placed on a magnetic stir bar drive motor, and the stir bar rotated rapidly enough to create a vortex at the surface of the solution, but not so vigorously as to cause droplets of the test chemical to separate from the pure chemical mass into the water portion of the partitioned solutions. It is important that the air space in the saturator be completely sealed from air exit or entry during use of the saturator. This air space will come into equilibrium with the water, and will not cause differences in solution concentrations with time. A pump is attached either to the dilution water input tube or the output tube which delivers test solution to the test system. Liquid-liquid saturators have been used successfully in exposures of aquatic organisms to a variety of potential EDCs, including di-*n*-butyl-, butylbenzyl, and di(2)-ethylhexyl phthalate (Call *et al.* 2001), 1,4-dinitrobenzene (Geiger *et al.* 1985), methoprene (Ankley *et al.* 1998b), and nonylphenol (Kahl *et al.* 1997).

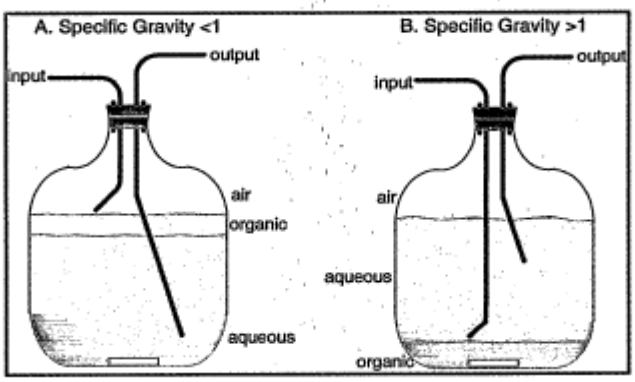

Figure 2. Liquid-liquid (slow-stir) saturation unit.

c. Glass Wool Column Saturator. This type of saturator has been used successfully with chemicals having water solubilities between 0.0074 and 104 mg/L (Kahl *et al.* 1999). Glass wool column saturators have been used successfully to generate stable water concentrations near solubility with chemicals that are liquids, but have specific gravities near 1, and do not maintain a discrete layer separate from the dilution water when mixed with a stir bar. This type of saturator can also be used to generate test concentrations near water solubility with chemicals that are solids at room temperature. The system is comprised of a clear glass tube approximately 2.5 cm in diameter and 1 m in length bent in a U-shape (primarily for economy of space) and packed with cleaned (solvent-rinsed) glass wool (Fig. 3). The test chemical is dissolved in a solvent (e.g., acetone, ethanol) which is added to the glass wool and drawn through the glass wool with a vacuum applied at one end. About one-half of the chemical solution should be added from each end of the glass tube to ensure uniform coating of the glass wool (for greater detail, see Kahl *et al.* 1999). When the column is completely vacated of the solvent, both ends of the tube are plugged with stoppers. Each stopper has a tube inserted and a pump is used to push water through the column. For some chemicals, a single pass of water through the column may result in the desired chemical concentration in the solution. Other chemicals may need several passes utilizing a recirculation system that pumps to a stock reservoir and back through the column. In the latter case, the concentrated stock is drawn from the stock reservoir. This type of stock solution-generating system has been used successfully with a variety of potential EDCs including DDT, DDD, and DDE (Hoke

et al. 1994), dieldrin (Hoke *et al.* 1995), methoxychlor and methyltestosterone (Ankley *et al.* 2001), polycyclic aromatic hydrocarbons (Erickson *et al.* 1999), and vinclozolin (Makynen *et al.* 2000).

The capacity of any of these saturators to maintain a constant concentration of chemical in solution will diminish with time of use. A general rule is to use at least twice the amount of chemical as the test may require to maintain a uniform concentration over the entire exposure period. A determination of the proper amount to "load" into the saturator can be calculated from information on water solubility of the test chemical (either from the literature or preliminary testing), volume of solutions in test chambers, number of dilutions, number of replicates per treatment, number of solution volume exchanges daily, and duration of the test. Multiple columns may be required over the course of one experiment.

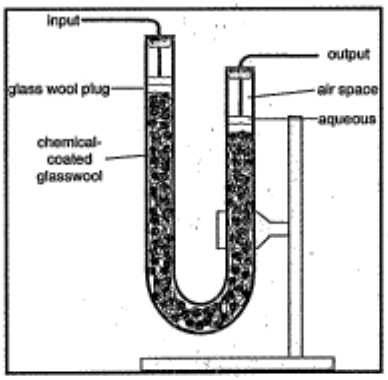

Figure 3. Glass wool column saturation unit.

d. Solvent Carriers. Certain chemicals may be insufficiently soluble in water or too unstable to use a saturator and generate a solution of sufficient concentration to conduct a test. It may be necessary to dissolve this type of chemical in a solvent carrier to achieve a test concentration near water solubility. Solvents may be injected into a dilution system for proportional dilution or directly injected into dilution water to prepare test solutions of desired chemical concentrations. In these experiments corresponding solvent-only control chambers are required.

7. Number of Treatments

The minimum number of treatments (test concentrations) is two (three if a concentration (dose)-response is sought), in addition to appropriate controls. Additional test concentrations are highly desirable but are, of course, dependent upon availability of resources. A single control treatment is required in the case of aqueous exposures. Two control treatments are required if chemical administration is via either solvent carrier, i.p. injection, or feeding (i.e., a solvent or sham-treated control in addition to a non-solvent or non-injected control).

8. Replication

The test chamber is considered to be the experimental unit or replicate. As the number of replicates per treatment increases, the number of degrees of freedom increases, the width of the confidence interval for point estimates decreases and, as a consequence, power of the

hypothesis test increases. The sensitivity of the test will be dependent upon the inherent variability of the endpoint under consideration as a function of the number of replicates employed in the treatments and control. Thus, every effort should be made to reduce among-sample variability of the measured endpoints. The greater the variance in measured effects between individuals within a replicate and the resultant means within a treatment, the greater the number of replicates needed to measure significance at any confidence level.

It is important that interpretation of the data facilitates evaluation of the potential for Type I and Type II errors. Type I error is the error of rejecting a null hypothesis (no difference between the control and treatment means) that is true; Type II error is the error of failing to reject the null hypothesis when it is false. For the purposes of a screening test, both types of error are critical, but the Type II error is more important because it could result in not detecting a possible EDC. Avoidance of the Type II error is accomplished by increasing the power ($1-\beta$) of the test. Many statistical tests with biological data are conducted with an α value of 0.05 and a β value of 0.20, where the probability of a Type I error is 5% and a Type II error is 20%. A more conservative Type II error probability of 10% ($\beta = 0.10$) may be desirable to increase sensitivity of the test. Type II error minimization (i.e., increased test power) requires either that the test be conducted with a greater number of replicates per treatment or that variability between treatment measurements be reduced (Fig. 4).

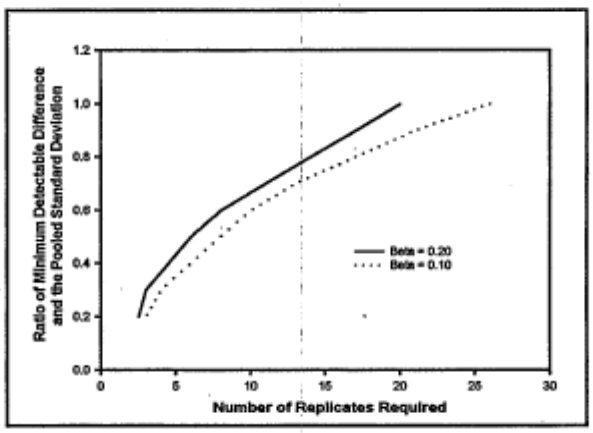

Figure 4. Number of replicates required per treatment to detect differences between treatment means with ANOVA using three treatments when $\alpha = 0.05$ and $\beta = 0.10$ or 0.20. Ratios are the desired minimum detectable difference in treatment means divided by the pooled standard deviation.

It is convenient to express the ratio of the difference between control/treatment means relative to their pooled standard deviation to demonstrate the impact of sample variability and sample size on detecting minimum differences (Fig. 4). For example, to demonstrate differences at the 95% confidence level ($\alpha = 0.05$) and with a power of 0.80 with four replicates per treatment, the variance between the means of the treatment(s) and control must be less than 50% of the difference between the means for which a significant difference may exist. If the variance is greater than 50%, more replicates will be required to detect differences. In the test report, power of the statistical analysis together with the confidence value should be reported.

For the purposes of this reproduction test, it is currently recommended that a minimum of four replicates (test chambers) be used per treatment. As with the number of concentrations tested, additional replication is desirable depending upon available resources. Further, as discussed above, depending upon established endpoint variability, replication may be increased in a study-specific manner. Unfortunately, because this test has not been used routinely, estimates of endpoint-specific variation are not yet robust enough for routine *a priori* design of tests with known power.

9. Performance Standard Test

A suggested approach for assessing the ability of a laboratory to conduct this assay is demonstration of expected test results with one or more known EDC(s). Conducting this type of test should reduce the frequency of the Type II error (i.e., having no effect when there is an expected effect) by demonstrating that the test system is performing reliably from both qualitative and quantitative (statistical) perspectives. Thus, it is recommended that each laboratory planning to perform this reproduction assay conduct a performance standard test with one or more known EDC(s). The results of this test should be reviewed and approved by knowledgeable scientific staff from outside of the laboratory prior to the initiation of routine screening assays. Chemicals potentially useful for this type of test might include strong estrogen receptor agonists such as β-estradiol or 17α-ethynylestradiol; or an aromatase inhibitor such as fadrozole (Kramer *et al.* 1998; Panter *et al.* 1998; 2002; Miles-Richardson *et al.* 1999a; Tyler *et al.* 1999; Korte *et al.* 2000; Ankley *et al.* 2002).

10. Reference Toxicant Tests

As a routine quality assurance option, it is desirable that reference toxicant tests be performed regularly (e.g., once annually), or concurrently with approximately 10% of other tests, to help verify aspects of general health of the test organism. This assay can be conducted using chemicals that are not necessarily EDCs, and can be comprised of short-term tests (e.g., 96-h) rather than the 21-d reproduction test. Some suggested reference toxicants include sodium chloride, potassium chloride, cadmium chloride, copper sulfate, sodium dodecyl sulfate, and potassium dichromate. If a reference test does not generate expected results, the test should be repeated immediately. Results from reproduction tests conducted during the interval of time since the last successful reference test are suspect. Reference toxicity tests must be conducted repeatedly with the same reference toxicant, at the same concentrations, in the same dilution water, using the same data analysis methods. A quality control chart is used to track the test results over time. The control chart should demonstrate that the test results are within ±2 standard deviations of the central tendency once several tests have been performed to establish the central tendency (Lewis *et al.* 1994). Reference toxicant tests could prove to be especially useful in situations where test animals are obtained in different lots from commercial vendors.

H. Procedure

1. Pre-Exposure Reproduction

   a. Selection of Fish. The pre-exposure phase of the test should be started with animals that have achieved reproductive maturity (typically a minimum of 120 to180 d old), as evidenced by initial development of secondary sex characteristics, but have not been held in a culture/test situation conducive to routine spawning. Ideally, these fish would be the offspring of several pairs of adults. Four females and two males should be randomly assigned to the replicate test chambers at each anticipated treatment concentration. Identification of gender may be difficult to resolve for some fish; these animals should not be used for the test. At this stage in development, males will exhibit nuptial tubercles, while females possess an ovipositor; in addition, males tend to be larger and darker than females from the same cohort (U.S. EPA 1987; Jensen *et al.* 2001). Some experience in working with this species will be required to accurately identify sex based on phenotypic characteristics, in particular, just prior to active spawning.

   In addition to the number of replicates (e.g., four) which actually will be exposed at each concentration, additional test chambers should be started; these can serve as "replacement" units for chambers where fish sex was misidentified, pre-exposure

spawning does not occur, and/or mortality is observed during the pre-exposure observation period.

b. Conditions. Pre-exposure reproduction observations must be performed using the same regime as described in Sections G.1, G.2, and H.2. The pre-exposure initiation of sexual reproduction must occur in the same chambers and under the same conditions (i.e., water characteristics, flow rate, lighting, temperature, feeding, and maintenance) as the 21-d reproduction test, to prevent unnecessary agitation of the fish at initiation of exposure to the test chemical. This monitoring phase establishes both the reproductive success of the spawning animals, and provides a quantitative chamber-specific baseline for potential statistical comparison after initiation of the chemical exposure. That is, each test chamber can serve as its own control.

c. Observations. Organisms in each test chamber should be observed daily for obvious alterations in secondary sex characteristics (i.e., nuptial tubercles, dorsal pad, and darkened coloration in males; and distended abdomen and swollen ovipositor in females), and reproductive behavior.

Observations should be made daily of spawning activity, and records maintained for the number of spawns and number of eggs per spawn. It also would be desirable to assess aspects of development of resultant embryos, such as fertility, embryo hatch, and gross appearance of newly-hatched larvae. If the animals are not reproducing

satisfactorily by 21 d, the pre-exposure portion should be terminated, and a thorough evaluation performed of the test system and fish to determine why satisfactory reproduction had not occurred.

d. Suitability for Testing. The pre-exposure period may last from 14 (minimum) to 21 (maximum) d. The exposure can be initiated once it has been established that regular spawning is occurring in each test chamber every 3 to 4 d. Based on information from tests conducted at MED to date (Ankley *et al.* 2001; 2002; Jensen *et al.* 2001; 2002), a minimal fecundity of 15 eggs/f/d would be expected for the test conditions described herein. The eggs should be successfully fertilized by the males, with a minimal fertility of 90%. Embryo hatch should be ≥90% and fry must be normal in appearance and swimming behavior; this is particularly critical when assessment of the F_1 generation is of interest. If these criteria of fecundity, fertility, embryo hatch, and fry appearance are not met, fish in the test chambers that fail to meet the criteria should not be used. If the criteria can not be met in all four of the test chambers during the pre-test conditioning, the fish should be discarded, the test system cleaned and evaluated, any necessary changes made, and a pre-exposure period initiated with a new group of fish.

2. Conditions of the 21-d Reproduction Test

a. Duration. The target exposure duration is 21 d. This exposure period should be sufficient for healthy females to produce several clutches of eggs. As such, this time period allows for a robust assessment of reproductive behavior, fecundity, fertility, and, if desired, embryo development and hatching success. In addition, the 21-d test period should help optimize exposure of the fish to relatively hydrophobic chemicals that require a period of time to reach steady-state concentrations in the animal. In the context of screening, instances may occur in which a test is terminated in less than 21 d. For example, it may be that the test chemical is clearly causing such significant adverse effects (e.g., lethality, cessation of spawning, observation of abnormal secondary sex characteristics, or behavior not consistent with reproduction, etc.) that it is clear that: (1) the test should be reinitiated at lower chemical concentrations, or (2) from a screening perspective, an effect consistent with an EDC has been established.

b. Loading. The total mass of fish in the test chambers must not affect the results of the test through alterations in water quality. Specifically, the total mass must be low enough to: (1) allow a ≥60% or greater dissolved oxygen saturation in the test solutions at all times (i.e., ≥4.9 mg/L at 25°C), (2) prevent concentrations of metabolic products (predominantly ammonia) from exceeding levels that affect animal behavior and health, and (3) prevent stress due to crowding. It is

recommended for a test of this type that the loading not exceed 0.5 g fish/L of solution passing through the test chamber per 24 h, or an instantaneous loading rate of 5 g fish/L (ASTM 2000a). The instantaneous loading rate for the test conditions suggested herein (i.e., four females and two males/10 L) is approximately 2 g fish/L.

c. Light and Temperature. Room light intensity of 100-foot candles during the lighted portion of the day is recommended. The light can be provided by fluorescent or incandescent lights. A photoperiod of 16 h light and 8 h dark, at a temperature of 25 ± 1°C is recommended. During a test, the water temperature should be monitored continuously in at least one test chamber, and daily in all others, and documented as being maintained within ±1°C.

d. Feeding. In tests where the route of test chemical administration is via either aqueous exposure or i.p. injection, the food should be frozen brine shrimp (*Artemia*) which have been thawed. The test animals should be fed the brine shrimp *ad libitum* twice daily. The frozen brine shrimp should be obtained from an established commercial source, and should have been demonstrated to result in acceptable survival, growth, and reproduction of the fathead minnow under culture conditions. Chemical characterization of the shrimp for common environmental contaminants that might affect endocrine function (e.g., organochlorine pesticides, PCBs) would be prudent. The amount of brine shrimp fed to each test chamber should be enough to nearly

satiate the fish with little, if any, food remaining at the bottom of the test chamber at the end of each feeding.

If test chemical administration is via the diet, the fish should be fed commercially available fish food with the chemical incorporated into it. One method of adding the chemical of concern to the food employs a rolling technique, in which a glass vessel containing the food and the chemical dissolved in an appropriate solvent (e.g., acetone, hexane) is mixed on a roller mill (Fernandez *et al.* 1998). Large lots of food can be made in a few minutes and, if storage life permits, refrigerated for use during the study. The chemical in the food should be measured for uniformity and stability during storage.

It should be noted that there has been little, if any, toxicological research in which fathead minnows have been dosed via incorporation of chemical into the diet. And, there is some uncertainty as to palatability/suitability of commercial pelleted feed to fathead minnows, particularly in longer-term (reproduction) studies. For these reasons, the dietary route, although conceptually reasonable is, at present, the least preferred manner of test chemical delivery.

e. Test Concentrations. The highest concentration used for the 21-d reproduction test should not have caused significant mortality in the range-finding assay (note that this may be at water solubility), and the lower concentration should normally be at a factor

of five to 10 times lower than the highest test concentration. The use of two test concentrations in this fashion not only enables at least some consideration of unexpected concentration (dose)-response relationships, but also provides a safety factor (lower concentration) should the high concentration prove to be lethal over the 21-d exposure. Alternatively, some studies may be undertaken using ≥ three test concentrations using a semi-log concentration series of 100, 32, 10, 3.2 and so forth (Panter *et al.* 1998). Concentrations of test chemical at each treatment level must be sufficiently uniform during the test to ensure that the fish are exposed to desired concentrations. In an aqueous exposure, concentrations at different levels should not overlap (e.g., due to cyclical trends associated with degradation). At a given exposure level, the measured concentration of the test chemical should not be less than 50% of the time-weighted mean concentration for more than 10% of the duration of the test (ASTM 2000b). In addition, the measured concentration should not be greater than 30% higher than the mean concentration for more than 5% of the duration of the test (ASTM 2000b). It should be recognized that once the exposure starts, there generally will be a gradual increase in the concentration during the first day. For example, with the recommended exchange rate of six volumes/d (42 ml/min) and a volume of 10 L per chamber, 95% replacement will not occur until approximately 12 h after test inititation (Sprague 1969). Analytical determinations of chemical concentrations should be made as frequently as possible with a minimum frequency of once weekly in at least one of the replicate test chambers at each concentration, and preferably in all chambers on the first and last days of the test. The variability of both the sampling

and the analytical procedures should be determined before initiating a test to ensure that observed variability in measured concentrations is not due solely to analytical techniques.

If the test chemical exposure route is by i.p. injection or the diet, tissue residue analyses should demonstrate a direct relationship between body burden and treatment. Further, it would be desirable to analyze the stock solution used either to spike the diet or for the i.p. injections to verify chemical concentrations prior to treatment of the animals.

The test chemical must be as pure as practical (>90% minimally, unless the chemical does not exist in this purity range).

f. Controls. Dependent upon route of chemical administration, one or more sets of controls will be required. If the chemical exposure route is aqueous, only a non-exposed control is required. If the chemical exposure route is aqueous, but with a solvent carrier, a solvent control is required in addition to a control that is not exposed to solvent. If the route of exposure is via i.p. injection, a sham- (e.g., corn oil) injected control is required in addition to a non-injected control. If the route of chemical exposure is via the diet, controls both without and with manipulated (i.e., spiked with solvent) food are required.

g. Fish. Organisms for the initiation of a laboratory culture should be obtained from a source which has a verified fathead minnow culture. Brood stock must be adapted to laboratory conditions and free of disease. Embryos make the best initial stock of fish because they are easiest to transport and most likely to be free of disease. Purchased fish need to be observed in the testing laboratory for a time suitable to ensure that they are in good health. Stock from wild populations should be avoided unless cultured through at least one generation to ensure they are disease-free and of adequate vigor (U.S. EPA 1987). Animals for starter cultures are available from the U.S. EPA Aquatic Biology Branch, Quality Assurance Research Division, EMSL-Cincinnati Newtown Facility, 3411 Church Street, Newtown, OH 45244 (Telephone: 513-533-8114); the U.S. EPA Mid-Continent Ecology Division, 6201 Congdon Boulevard, Duluth, MN 55804 (Telephone: 218-529-5000); or any of several commercial suppliers with species-verified stocks, such as Environmental Consulting and Testing, 1423 North Eighth Street, Superior, WI 54880 (Telephone: 715-392-6635 or 800-377-3657; web address: www.ectesting.com); or Aquatic Research Organisms, Hampton, NH 03842 (Telephone: 603-926-1650; e-mail: arofish@aol.com; web address: www.holidayjunction.com/aro).

For laboratories not intending to start a fathead minnow culture, animals of a known age (prior to maturity) may be obtained from commercial suppliers for testing. If this is done, animals should be held for at least 1 month prior to initiating the pre-exposure period for the test.

h. Cleaning of Test Chambers. All components of the test apparatus that will have contact with water or chemical must have been adequately cleaned with detergent, solvents, and deionized water to prevent the occurrence of disease organisms or chemical residues from any previous use of the apparatus. During the pre-exposure period and the test, chambers need regular cleaning to remove biological debris (excess food, excrement). This can be effectively accomplished with a small (0.8-1.0 cm i.d.) Tygon® siphon tube. After one week, algal and fungal growth may begin to appear, interfering with observations of fish, reducing the adherence of eggs to the spawning substrates, and potentially reducing bioavailability of the test chemical. The spawning substrates should be removed from the test chamber and scrubbed. The growth on the inside of the test chamber can be removed by carefully (to avoid agitating the fish) scraping the inside walls with a sponge or razor blade inserted into a long-handled holder. Once the material scraped from the chamber walls has settled to the bottom, it can be removed with a siphon tube, and the spawning substrates replaced. It is essential to treat all test and control chambers identically to avoid introducing bias.

3. Test Chemical Administration

a. Flow-Through Aqueous-Delivery. Test concentrations of the chemical can be achieved by blending known concentrations of stock solution with dilution water in

mixing cells of the exposure apparatus, from which solutions of desired concentrations are allocated to test chambers. To achieve comparable flow rates of water among the treatment groups, separate reservoirs containing different concentrations of stock solution should be prepared for each treatment level, and the pumping rate from each reservoir should be equivalent. The stock solutions may be prepared as specified previously (Section G.6), depending upon characteristics of the test chemical. Flow rates to the exposure chambers should deliver a minimum of six volume exchanges daily.

b. Flow-Through Aqueous/Solvent-Delivery. For tests in which a solvent or dispersant is used to facilitate delivery of the test chemical, a solvent or dispersant control must be included in the experimental design. This control should be identical in all aspects to the non-solvent control, except that it should contain the solvent or dispersant at the maximum concentration used in the different test chemical treatment levels. The delivery system should be maintained as described above for the water-delivery system. It may be necessary to increase the frequency of cleaning the delivery system and test chambers, due to the promotion of microbial growth by the solvent. Flow rates to the exposure chambers should deliver a minimum of six volume exchanges daily.

c. Intraperitoneal (i.p.) Injection. Certain chemicals may be so insoluble in water that aqueous exposures are not feasible. Alternatively, availability of the test material

may be so limited that a flow-through aqueous exposure is not feasible. In such cases, exposure via the diet or by i.p. injection may be required. When the route of chemical administration is via i.p. injection, two types of controls must be employed, a non-injected control and a sham-injected control. The test chemical should be dissolved in a non-toxic carrier, such as corn oil or a mixture of ethanol/corn oil (Korte *et al.* 2000; Kahl *et al.* 2001). The injection procedure involves first lightly anaesthetizing the fish with MS-222 (100 mg/L, buffered with 200 mg $NaHCO_3$/L), then carefully holding the fish in one hand in an inverted position, and inserting a sterile needle containing the desired volume (and concentration) of test solution just through the abdominal body wall (Fig. 5). The maximum needle size should be 26 gauge (ca. 1 cm in length), and no more than 10 µl of solution/g body weight should be injected in an individual fish. The needle should be inserted just anterior to the vent with the needle tip oriented anteriorly and at an angle of approximately 30 to 45° from the horizontal position. The orifice at the tip of the needle must be completely inserted through the body wall without entering internal organs to ensure that the entire volume desired is injected into the fish. After injection, the needle tip is withdrawn and the point of injection carefully examined to determine that there has not been any loss of solution. The injection process should be accomplished quickly (e.g., within 30 s), after which the fish should be immediately returned to its respective test chamber. Injected fish should be observed over the next several hours to ensure that they have recovered from the anaesthetic and have not suffered any ill effects from the injection. Depending upon the kinetics of metabolism/depuration of

the test material, multiple injections may be required during the 21-d test. Data from the MED laboratory indicate that i.p. injections at a frequency of up to one/week should not affect the reproductive and endocrinological endpoints associated with the test (Kahl *et al.* 2001).

When using the i.p. exposure route, flow rates (of clean water) to exposure chambers should deliver a minimum of six volume exchanges daily.

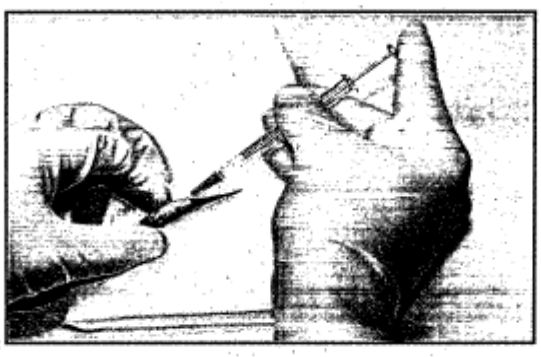

Figure 5. Handling technique for administering test chemicals to fathead minnows by i.p. injection.

d. Dietary Exposure. In cases where a dietary route of chemical administration is used, commercial fish feeds are recommended. One procedure that has been described for preparing a diet containing a highly insoluble chemical (2,3,7,8-tetrachlorodibenzo-*p*-dioxin) utilized an apparatus that delivered the test chemical dissolved in *n*-hexane to

the food pellets in a modified glass jar on a roller mill (Fernandez *et al.* 1998). The spiked hexane is removed from the food using a vacuum, and condensed and collected in a cold trap. It is recommended that a procedure similar to this be used to minimize exposure of the fish to the carrier solvent and, at the same time, achieve an acceptable degree of homogeneity of test chemical in the food. Two types of controls are needed when conducting dietary exposures. One should receive equal amounts of food that has not received any treatment of either solvent carrier or test chemical, and the second control should receive food that has been mixed with solvent carrier only. Food should be administered at a rate of about 2.5% of body weight per day. This should be adjusted so that no food remains in the test chambers at the end of each day.

Selection of target chemical concentrations for the food when using a dietary route may be challenging in that it is necessary to identify not only a target dose in the tissue of the organism, but some estimate of the likely dietary assimilation of the test chemical of concern. These estimates must be done relative to the actual rate of food consumption by the fish.

When using the dietary exposure route, flow rates (of clean water) to the exposure chamber should deliver a minimum of six volume exchanges daily.

4. Analytical Determinations and Measurements

a. Test Chemical. For tests with an aqueous route of administration, concentration of the stock solution must be measured initially and at least once weekly. In addition, the test chemical should be measured in the exposure water at least once weekly during the 21-d exposure. In the range-finding tests, the behavior of the chemical in solution also should be determined; this is a good opportunity to assess, under simulated test conditions, performance of the stock solution generation/delivery system. For chemicals that do not require range-finding tests, performance of the exposure system should be evaluated independently prior to testing. During the 21-d reproduction test, measurements should be made in at least one replicate test chamber of all exposure concentrations on each sampling day. The analytical procedures should follow U.S. EPA-approved or other validated standard methods, if available. When appropriate, reagent blanks, recoveries, and standards should be included whenever samples are analyzed. Balances and analytical instruments should have properly updated records of calibration or service, as appropriate. Standard reference materials should be analyzed, when available, to document accuracy of analyses.

Recovery of the test chemical from the exposure medium should be determined from a sufficient number of samples to allow an assessment of accuracy of the analytical process. A minimum of three samples is required; however, six or more analyses are preferred. Precision of the analyses is determined by collecting duplicate samples at a

frequency of at least 10% of the total samples analyzed. A relative percent difference (RPD) is calculated for each set of duplicate analyses, and an overall mean and standard deviation calculated for the entire number of duplicates analyzed. The method detection and quantitation limits for test chemicals must be defined.

b. Water Characteristics. Water quality characteristics must be measured regularly to document that they were sufficiently uniform to ensure that adverse biological effects due to water quality did not affect the outcome of the test. Temperature must be monitored daily and maintained at 25 ± 1°C. Dissolved oxygen in the test solutions must also be monitored daily to ensure a minimum concentration of 4.9 mg/L (60% saturation at 25°C). Alkalinity, hardness, pH, and un-ionized ammonia, as a minimum, should be measured at least weekly during the test in a high and low test concentration. In addition, since particulates and organic carbon may influence the bioavailability of many test substances, concentrations of both total organic carbon and particulate matter should be determined and documented as ≤5.0 mg/L (ASTM 2000b). Table 2 summarizes recommended ranges of values for water quality characteristics. Sample data forms, such as the example provided (Appendix A, Form A-1), should be prepared in advance and completed on a daily basis.

5. Measurement of Test Endpoints

Observations of the test animals (described below) should be made daily and/or at conclusion of the test.

a. Survival of Adults. Daily assessment of survival is necessary to provide a basis for expression and interpretation of reproductive output, that is, number of eggs per female per day. Unless unacceptable water quality excursions and/or disease occur, it is rare to observe mortality in untreated control animals. In animals exposed to the test chemical, overt lethality may occur, particularly in later portions of the assay not reflective of the initial (shorter) range-finding test. Mortality should be recorded daily on a form such as the example provided (Appendix A, Form A-1),

b. Behavior of Adults. Any abnormal behavior (relative to control animals) observed during the course of the chemical exposure should be monitored. This might include signs of general toxicity, including hyperventilation, uncoordinated swimming, atypical quiescence, loss of equilibrium, or abnormal feeding. From the standpoint of EDC screening, alterations in reproductive behavior, particularly loss of territorial aggressiveness by males, may be affected by EDCs (Kime 1998). Detailed descriptions of normal behaviors and the abnormal behaviors should be recorded to compare the changes that may be observed during the test. Because of the relative

subjectivity of this endpoint, it may be necessary to document behavioral alterations via photograph or videotape.

c. Fecundity. Fecundity, or total egg production, should be determined on the basis of surviving females per reproductive day for each test chamber. For example, if all four females survived the treatment in a given chamber in a 21-d exposure, there would be 84 female reproductive days. If one or more females dies during the course of the exposure, the number of female reproductive days would be reduced accordingly in a "pro-rated" fashion.

Fathead minnows usually spawn in the early morning hours so they should not be disturbed during this time except for a morning feeding. The spawning substrates should be checked for the presence of eggs in the late morning. This period of time allows the opportunity for spawning to occur and the eggs to water harden. It is necessary to remove the substrates from the test chambers to prevent the eggs from being eaten by the adults, and for purposes of determining egg count and fertility. If no eggs are present, the spawning substrate is left in the test chamber; new substrates should be added to replace any that were removed. Data forms to record fecundity, such as the example provided (Appendix A, Form A-2), should be prepared in advance and completed on a daily basis.

d. Fertility. After the spawning substrate has been removed from the test chamber, the eggs should be carefully rolled off the substrate with a gentle circular motion of an index finger (Gast and Brungs 1973) and visually inspected under appropriate magnification. If more than one distinct stage is present, consider each stage as one spawning and handle separately as described below. If spawning occurred that morning, embryos will typically be undergoing late cleavage, and determination of fertility (number of embryos/number of eggs x 100) is easily achieved. Occasionally, spawning occurs in the afternoon and embryos will be close to tailbud stage by the following morning. A detailed embryonic staging sequence for fathead minnows is provided in U.S. EPA (1996b). Infertile eggs are opaque or clear with a white dot where the yolk has precipitated; viable embryos will be clear for the first 36 to 48 h until they reach the eyed stage. An alternative to the microscopic approach to determining fertility is to enumerate eyed embryos at this time. If the latter technique is used, the embryos need to be held in a system apart from the adults in a water bath with aeration (U.S. EPA 1987) or placed in incubation cups (U.S. EPA 1982). The fertility rate in control animals generally will exceed 95%.

e. Embryo Hatch (Optional Endpoint). Regardless of the method used to determine fertility, if information concerning hatching success and/or subsequent development is desired, the embryos need to be maintained in incubation chambers. Depending on study objectives, embryos can be held in clean water or water containing the same concentration of chemical to which the parents were exposed to continue an F_1

generation exposure. Incubation cups can be made from glass cylinders or jars with the bottoms cut off and nylon or stainless steel screen glued to the bottom with silicone adhesive (U.S. EPA 1982). The cups can be suspended in a tank with an oscillating water level to ensure that the embryos are always covered and that water regularly flows into and out of the cups without creating excessive turbulence. Alternatively, they can be held in incubation cups containing clean water under static conditions. If this is done, the incubation cups should be held in a constant-temperature water bath and water inside the cup renewed daily. Discrete spawns should be screened with a dissecting scope to remove empty or opaque shells, or abnormal embryos, and a subsample (e.g., 50) should be selected at random and placed in the incubation cup. At 25°C, untreated animals will hatch in approximately 4 to 5 d. Embryos should be inspected daily and any dead animals counted and removed. A small container should be placed under the incubation cup prior to removal from the water bath for examination under a dissecting scope. Potential endpoints include time to complete hatch, total number of embryos hatched, and number of normal larvae at hatch. The hatching rate of control animals typically is in the range of 95 to 98%.

f. Larval Survival and Morphology. (Optional Endpoint). Newly hatched fry should be observed for general physical appearance and behavior. Healthy fry should actively swim about the incubation cup. Gross morphological abnormalities, described by Sharp (1991), which may be observed include lordosis, scoliosis, kyphosis, retarded

swim bladder development, edema, and craniofacial abnormalities. Survival of the larvae can easily be assessed through yolk sac absorption (~1 to 2 d post-hatch at 25°C); if estimates of survival are required after this, the animals must be fed (generally live *Artemia* nauplii; U.S. EPA 1987). An extended observation period may be necessary to detect chemicals that cause delayed mortality, such as Ah receptor agonists (Elonen *et al.*, 1998). Fry are fed live brine shrimp nauplii (< 24 h old) until approximately 30 d of age (U.S. EPA 1987), possibly supplemented from 20 d of age with powdered dry food (Länge *et al.* 2001).

g. Secondary Sex Characteristics. Observations of physical appearance of the adults should be made over the course of the test and at conclusion of the study. From the perspective of screening EDCs, characteristics of particular importance include body color (i.e., light or dark), coloration patterns (i.e., presence or absence of vertical bands), body shape (i.e., shape of head and pectoral region, distension of abdomen), and specialized secondary sex characteristics (i.e., number and size of nuptial tubercles, size of the dorsal pad and ovipositor). Notably, chemicals with certain MOA can cause the abnormal occurrence of certain secondary sex characteristics in the opposite sex; for example, androgen receptor agonists, such as methyltestosterone, can cause female fathead minnows to develop nuptial tubercles (Smith 1974; Ankley *et al.* 2001), while some estrogen receptor agonists may decrease number or size of nuptial tubercles in males (Miles-Richardson *et al.* 1999a; Harries *et al.*, 2000; Länge *et al.* 2001). Any abnormalities in appearance should be documented and, to the

extent possible, quantified (e.g., counting tubercles). It may be advantageous to photograph or videotape the appearance of the fish near the end of the test. Because some aspects of appearance (primarily color) can change quickly with handling, it is important that qualitative observations be made prior to removal of animals from the test system. This type of qualitative assessment might be enhanced through the use of photographs or videotape. Other endpoints, such as the number and size of nuptial tubercles, can be quantified directly or in preserved specimens.

An example of a data form to record all pertinent sample data at the conclusion of the test is provided (Appendix A, Form A-3).

h. Blood Collection. At conclusion of the 21-d test, the fish should be netted from the exposure chambers one at a time and anaesthetized with MS-222 (100 mg/L, buffered with 200 mg $NaHCO_3$/L). After anaesthetization, the caudal peduncle should be partially severed with a scalpel blade and blood collected from the caudal vein/artery with a heparinized microhematocrit capillary tube (Figure 6). After the blood has been collected, the plasma is quickly isolated by centrifugation for 3 min at 15,000 *g*. If desired, percent hematocrit can be determined following centrifugation. The plasma portion is then removed from the microhematocrit tube and stored in a centrifuge tube with 0.13 units of aprotinin (a protease inhibitor) at -80°C until determination of vitellogenin and sex steroid concentrations can be made. Depending

on size of the fathead minnow (which is sex-dependent), collectable plasma volumes generally range from 20 to 60 µl per fish (Jensen *et al.* 2001).

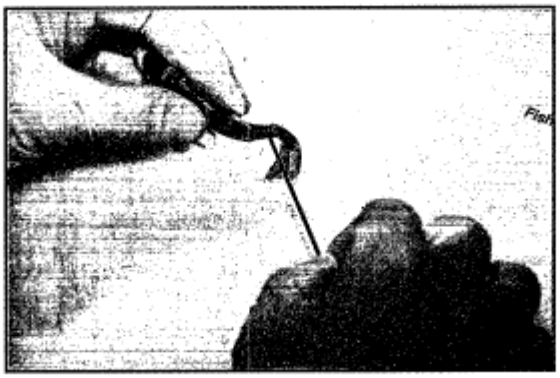

Figure 6. Fathead minnow blood collection process.

i. Vitellogenin. Vitellogenin is a phospholipoglycoprotein precursor to egg yolk protein that is synthesized in the liver of sexually-mature females of all oviparous species; the production of vitellogenin is controlled by interaction of estrogens, predominantly β-estradiol, with the estrogen receptor (Sumpter and Jobling 1995; Kime 1998). Significantly, males maintain the capacity to produce vitellogenin in response to stimulation with estrogen receptor agonists; as such, induction of vitellogenin in males has been successfully exploited as a biomarker specific for estrogenic compounds in a variety of fish species, including the fathead minnow (Kramer *et al.*

1998; Panter *et al.* 1998; 2002; Parks *et al.* 1999; Tyler *et al.* 1999; Harries *et al.* 2000; Korte *et al.* 2000; Ankley *et al.* 2001; Länge *et al.* 2001).

Different methods are available to assess vitellogenin production in fish; a measurement technique that is both relatively sensitive and specific is determination of protein concentrations in plasma via enzyme-linked immunosorbant assay (ELISA). In performing quantitative ELISAs with fathead minnow samples, some have utilized polyclonal or monoclonal vitellogenin antibodies prepared using protein from other fish species that cross-reacts with fathead minnow vitellogenin (Tyler *et al.* 1999; Harries *et al.* 2000). Others have used fathead minnow-specific polyclonal vitellogenin antibody and purified fathead minnow vitellogenin protein for the ELISA (Parks *et al.* 1999; Korte *et al.* 2000). Using the fathead minnow-specific approach, detailed instructions for the measurement of fathead minnow vitellogenin are given in Appendix B.

j. Sex steroids. Plasma concentrations of β-estradiol, testosterone, and 11-ketotestosterone can be determined using radioimmunoassay (RIA) techniques optimized for the relatively small sample volumes obtained from the fathead minnow (Jensen *et al.* 2001). There is emerging literature documenting concentrations of these sex steroids in the fathead minnow under both normal conditions and after treatment with various EDCs (Giesy *et al.* 2000; Makynen *et al.* 2000; Ankley *et al.* 2001; 2002; Jensen *et al.* 2001; 2002). In mature females, both β-estradiol and

testosterone are easily measured, while 11-ketotestosterone concentrations typically are low to non-detectable. In mature males, concentrations of 11-ketotestosterone are two- to five-fold greater than those of testosterone, and β-estradiol concentrations are low to non-detectable (Jensen *et al.* 2001).

Detailed instructions for the measurement of fathead minnow plasma sex steroids using RIA are given in Appendix C.

k. Gonadosomatic Index. An assessment of reproductive status can be determined by measurement of the gonadosomatic index (GSI). The GSI is the weight of the ovaries or testes relative to the total body weight of the fish (GSI = 100 x gonad weight (g)/body weight (g); Crim and Glebe 1990). After removing blood, wet weight should be measured on individual fish to the nearest 0.01 g. A ventral incision through the body wall from the anus to the isthmus should be made and both ovaries in females and both testes in males carefully removed with fine forceps and placed into a pre-weighed weighing pan. Gonad wet weight should be measured to the nearest 0.1 mg.

Typical GSI values for the fathead minnow range from 1 to 2% for males and 8 to 13% for females (Smith 1978; Jensen *et al.* 2001). In fractional spawning fish, like the fathead minnow, the ovaries undergo rapid cyclical changes over relatively short periods of time as successive batches of eggs are produced. Female GSI varies

significantly as a function of spawning interval, with the smallest values occurring just after spawning followed by an increase of 45% within 3 to 4 d, just prior to the next spawning event (Jensen *et al.* 2001). This means that the ovaries in breeding females can vary considerably between individuals within treatments at any one point in time. This could make identifying alterations in GSI as a result of exposure to EDCs difficult if the normal variation associated with the reproductive cycle exceeds effects associated with exposure to the test chemical. Nevertheless, significant effects on GSI have been reported in fathead minnows following exposure to various EDCs (Panter *et al.* 1998; Makynen *et al.* 2000; Ankley *et al.* 2001; 2002).

l. Gonadal histology. Histological evaluation of the gonads is an important component of the test because it may reveal specific alterations in the gonads leading to possible impacts on reproductive output and also provide possible insight into the mechanism of action of potential EDCs. Several recent studies involving exposure of fathead minnows to EDCs have included an examination of gonadal histology. Most, but not all, of these studies have considered effects on both ovaries and testes. Ovaries of females exposed to strong estrogen receptor agonists such as β-estradiol or ethinylestradiol exhibit fewer mature, and more atretic, follicles (Miles-Richardson *et al.* 1999a; Länge *et al.* 2001). Kramer *et al.* (1998) and Miles-Richardson *et al.* (1999a) noted that β-estradiol exposure resulted in a sustained increase in plasma vitellogenin in females. They suggested that sustained abnormally high vitellogenin levels interfere with final maturation and release of oocytes from the ovary, possibly

by inhibiting gonadotropin release by the pituitary. Testicular effects varied, depending on concentration, from slight degeneration, especially involving spermatozoa, to clear atrophy and appearance of ova-testes. Moreover, Miles-Richardson *et al.* (1999a) reported that a concentration-dependent proliferation of Sertoli cells containing remains of spermatozoa accompanied degeneration of spermatozoa. Exposure to methoxychlor, a weak estrogen receptor agonist, resulted in increased follicular atresia (Ankley *et al.* 2001), a finding consistent with a substantial reduction in fecundity during a 21-d exposure. Exposure to nonylphenol, another weak estrogen receptor agonist, resulted in significant necrosis of germ cells and spermatozoa accompanied by hyperplasia and hypertrophy of Sertoli cells containing germ cell remnants (Miles-Richardson *et al.* 1999b). Examination of the ovaries of females exposed to the strong androgen receptor agonist, methyltestosterone, revealed no postovulatory follicles (corpora lutea) (Ankley *et al.* 2001). Instead, maturation of younger follicles was suppressed and there were numerous pre-ovulatory atretic follicles, consistent with the observation that spawning ceased immediately following exposure to the androgen. Methyltestosterone-exposed testes differed from controls in that they appeared to be stimulated to near exhaustion of germinal epithelial stages. The germinal epithelium was much thinner, and spermatogenic activity more scattered than in control testes (Ankley *et al.* 2001). The effects of two putative anti-androgens also have been evaluated. Makynen *et al.* (2000) observed significantly reduced oocyte diameters following exposure to vinclozolin. That is, the ovaries exhibited retarded maturation,

especially regarding yolk deposition in oocytes. Similarly, exposure to flutamide resulted in more early-stage follicles and an increase in atretic follicles (Jensen *et al.* 2002). Both of these factors could indicate decreased egg production via a delay in oocyte maturation resulting in fewer eggs being produced in a given time or a greater than average number of follicles undergoing resorption before reaching maturity. Exposure to the aromatase inhibitor, fadrozole, resulted in a near (lowest concentration) or complete (higher concentrations) absence of corpora lutea, increased numbers of pre-ovulatory atretic follicles, and regression to an earlier developmental stage (higher concentrations) (Ankley *et al.* 2002). Judging by histological criteria, fadrozole appeared to reduce fecundity by preventing maturation of oocytes.

Routine histological procedures can be used to assess fathead minnow gonads. Appendix D discusses two approaches to fixation and embedding appropriate for fathead minnow gonadal histology: a traditional paraffin-based approach and a more modern methacrylate-based histological procedure. Either technique is acceptable in the context of the test described in the document. An assessment of normal gonadal histology and routine methods for evaluating EDC-induced histological changes in reproductively-mature fathead minnow gonads is given in Appendix E.

m. Tissue Residues (Optional Endpoint). Adult tissue residue analysis verifies that the chemical entered the fish, and also offers an opportunity to calculate chemical-

specific bioconcentration or bioaccumulation factors. Further, the relationship of specific tissue residues to toxic effects can be valuable in terms of across-chemical, across-species, and laboratory-to-field extrapolations (Jarvinen and Ankley 1999). Because of potential among-fish variations in tissue residues, individual fish should be analyzed whenever possible. If it is necessary to pool samples to obtain adequate tissue mass for analyses, it is recommended that males and females from any given treatment be kept separate because of possible gender-related differences in bioconcentration (Makynen *et al.* 2000; Ankley *et al.* 2001). For some types of test chemicals (e.g., nonionic organics), measurement of organism lipid content could prove useful for assessing bioconcentration/bioaccumulation data (Ankley *et al.* 2001).

n. Other Endpoints (Optional). Cytochrome P450 aromatase (CYP19) catalyzes a key step in the conversion of C19 androgens to C18 estrogens. The ovarian form of this enzyme is the major source of circulating β-estradiol in females (Mommsen and Walsh 1988). In fish, both males and females also have significant aromatase activity in brain tissue (Gelinas *et al.* 1998; Melo *et al.* 1999). Inhibition of this critical enzyme could result in reduced levels of β-estradiol and alter the expression of many genes controlled by the estrogen receptor (e.g., vitellogenin production).

Activity of the enzyme can be measured using a simple assay in which radiolabeled androstenedione releases tritiated water during aromatization and conversion to

estrone (Thompson and Siiteri 1974). Several methods have been reported in the literature for measuring this activity in fish tissue (Pelissero *et al.* 1996; Chang *et al.* 1999; Melo *et al.* 1999; Shilling *et al.* 1999; Kitano *et al.* 2000). Because of the small size of the fathead minnow, a modification of the microassay of Melo *et al.* (1999) has been used with brain tissue from fathead minnows (Ankley *et al.* 2002). In the latter study, adults exposed to the aromatase inhibitor, fadrozole, exhibited a concentration-dependent decrease in brain aromatase activity.

The expression of the aromatase gene has also been determined by measuring gonadal aromatase mRNA (Kitano *et al.* 2000; Scholz and Gutzeit 2000). Determination of ovarian aromatase mRNA levels is also possible in individual fathead minnows. Using the polymerase chain reaction (PCR) and consensus primers designed from the DNA sequence of aromatase genes in other fish, an amplified segment of the fathead minnow aromatase cDNA was obtained (GenBank accession number AF288755). This segment of DNA was used to generate a biotin-labeled antisense RNA probe for use in measuring mRNA levels via a ribonuclease protection assay (RPA) (J. Korte, unpublished data). The same sequence information could be used in a quantitative real-time PCR method. Measurement of aromatase expression via mRNA levels is particularly useful in fathead minnow ovaries since the small size of the tissue makes direct enzyme activity measurement challenging in individual fish. It is not currently known if changes brought about by aromatase inhibitors, such as fadrozole, will alter the level of expression of the gene at the mRNA level.

Since the effects of androgens must be manifested through binding to the androgen receptor, some means of measuring changes in the level of the receptor could be effective in identifying EDCs. Plasma β-estradiol levels are known to be capable of up-regulating the level of the estrogen receptor in fish liver (Pakdel *et al.* 1991; Andreassen and Korsgaard 2000). Although regulation of the androgen receptor is not well understood, there is some evidence of tissue-specific up- and down-regulation by 11-ketotestosterone (Todo *et al.* 1999). Observations with female fathead minnows have shown the appearance of nuptial tubercles when exposed to the androgen receptor agonist, methyltestosterone (Ankley *et al.* 2001). These tubercles are normally only found in sexually-mature males. Since the appearance of tubercles would seem to be a response mediated through the androgen receptor, it is possible that the appearance of male secondary sex characteristics in female fish involves changes in the level of the androgen receptor. Using PCR and consensus primers designed from the DNA sequence of androgen receptor genes in other fish, an amplified segment of fathead minnow androgen receptor cDNA was obtained (GenBank accession number AJ277866). This information was used to prepare biotin-labeled antisense RNA probes for use in a RPA to measure androgen receptor mRNA (J. Korte, unpublished data). Conversion to a quantitative real-time PCR method could also be easily accomplished.

I. Treatment of Results

1. Overview

The reproduction test generates quantitative data on adult survival, GSI, plasma concentrations of vitellogenin and sex steroids, fecundity, fertility, and, if desired, embryo hatch and larval survival. The null hypothesis that no differences exist between the control fish and those exposed at the different treatment levels should be statistically evaluated for each test endpoint. Both types of statistical error (i.e., Types I and II) should be clearly specified in the presentation of the statistical results.

Time-dependent data (i.e., fecundity and fertility) potentially require the most complex analysis because this information is collected daily both before and after exposure to the test chemical. These data can be analyzed in many ways. Some methods are more appropriate than others for this type of data, which can be highly variable especially when the mean fecundity is large, temporally correlated, or skewed. Mean fecundity per period (i.e., pre-exposure or post-exposure) and test chamber can be relatively variable, and variance usually increases with mean fecundity. Also, the correlation between pre-exposure and post-exposure mean fecundity from the same test chamber may be low. The recommended analysis is based on a summary statistic for each test chamber, either the mean post-exposure fecundity or the ratio of the post-exposure to pre-exposure fecundity.

Data on other endpoints do not present the complexities of the time-dependent data (i.e., fecundity, fertility); they are measured only once at the termination of the experiment (adult survival, GSI, and plasma vitellogenin and sex steroid concentrations).

The use of the suggested statistical methods for routine data analysis does not require the assistance of a statistician. However, interpretation of the results may become problematic because of the inherent variability and occasionally unavoidable anomalies in the data. If the data appear unusual in any way, or fail to meet the necessary assumptions, a statistician should be consulted. Analysts who are not proficient in statistics are strongly advised to seek the assistance of a statistician before selecting the method of analysis and interpreting the results.

2. Analysis of Fecundity Data[1]

a. Data preparation and adjustment for mortality. The experimental period is divided into three time phases. The first seven days should be considered an acclimation period; omitting these data from subsequent analyses may reduce variability. The remaining period until the treatment is imposed is considered the pre-exposure period. Observations from day 1 of exposure until the end of the experiment are the post-exposure period.

---

[1] Although this section discusses fecundity data specifically, most of the considerations and approaches are also germane to fertility and, if collected, hatch data.

The mean (± SD) daily fecundity during the pre-exposure and post-exposure periods should be calculated for each test chamber. The daily observations of fecundity can be highly variable; calculating the mean fecundity reduces some of this temporal variability. Also, the experimental unit is the test chamber and analysis of the data is much simpler when summarized into one value per chamber. That single value per chamber could be the post-exposure fecundity, the ratio of fecundities, or some other function of the fecundity. The choice should be specified before the analysis is conducted.

When there is no mortality, the mean daily fecundity is an appropriate measure of spawning success. These experiments are designed to estimate sub-lethal effects so the default assumption is that test concentrations are low enough so that there should be no treatment-related mortality. However, random deaths may occur, and loss of one or more females may decrease the mean daily fecundity in that chamber. Two approaches could be used to adjust for random mortality.

1) Compute the mean fecundity on a per live-female per day basis. Compute the total number of eggs produced during the pre-and post-exposure periods. Then, compute the total number of female reproductive days during each period. The corrected mean is the ratio of the total number of eggs to the number of female reproductive days.

2) Omit fecundity data from any test chamber with mortality.

The first approach uses data from all chambers, but assumes that the mortality of a female has no effect on any other fish, that the mortality was not caused by exposure to the test chemical, and that there was no period of reduced fecundity prior to death. The second approach only assumes that the mortality was not caused by the test chemical, but it results in a smaller sample size and smaller statistical power. All the available biological information should be used to choose the more appropriate adjustment. In instances where the death of a male fish occurs, the second approach may be more appropriate.

The fecundity data should be plotted, both as a preliminary step to help detect anomalies and unsuspected trends or patterns in the responses, and as an aid in the interpretation of the results. These plots can also be used to visually check the validity of some assumptions that are necessary for several statistical analyses that may be used.

b. Choice of analysis. The summary statistics, one for each test chamber in the experiment, can be analyzed by many different statistical procedures. One approach is to use general characteristics of the data to identify an approach and possible alternatives. All tests are conducted on the same organisms, in a similar manner, and measure the same response, mean fecundity. The mean fecundity will not be the

same in each experiment, because different chemicals have different effects. However, every one of the tests provide information about the characteristics of the data that influence the choice of analysis.

Consideration of possible approaches for analyses can be organized into four decisions; recommendations for each decision should be based on general characteristics of the data. The recommendations below are tentative. They may change as this test is conducted more frequently, and general characteristics of the resultant data become better known.

1) Choice of response: It is possible to use mean fecundity during the post-exposure period or use the ratio of the post-exposure and pre-exposure fecundity. A third option is to use the difference between post-exposure and pre-exposure mean fecundity. The choice of response affects the statistical power of the analysis. If the correlation between pre-exposure and post-exposure fecundity is high or moderately high, analysis of the ratio is more powerful. If the correlation is low, analysis of the post-exposure data alone is more powerful. Often, because of the large inherent variability, the observed correlation will be relatively small, so it may be most common to analyze the post-exposure mean.

2) Choice of hypothesis to be tested: It is possible to test the hypothesis of equal means, or test for a monotonic concentration (or dose) response, or test whether

specific treatments differ from the control. The traditional hypothesis, tested by analysis of variance (ANOVA), is equal means. This is appropriate for any set of treatments, but it ignores extra information available when the treatments are a sequence of test concentrations. If the concentration (dose)-response curve is expected to be monotonic, a more powerful test is that of ordered alternatives. This tests the hypothesis that $\mu_C \geq \mu_L \geq \mu_H$, where μ_i is the mean fecundity in the ordered treatments, ranging from control to higher concentrations. If desired, this test can be followed by comparisons of specific concentrations to the control using an adjustment for multiple comparisons.

3) Choice of parametric (ANOVA) or nonparametric analysis: This choice affects the statistical power and validity of the α level. It is also influenced by the ease of computation for a test. If the data satisfy the ANOVA assumptions (data are independent, normally distributed, and have equal variances), ANOVA is slightly more powerful than nonparametric alternatives. When the experimental design is balanced, i.e. equal number of replicates of all treatments, violations of the normality and equal variance assumptions have little effect on the α level, but decrease the statistical power. Nonparametric ordered alternatives tests require moderately complicated calculations and are not available in commonly-used computer software. A good nonparametric test of ordered alternatives, the Jonckherre-Terpstra test, is available in SAS (SAS Institute, Cary, NC) or SPSS (SPSS, Inc., Chicago, IL); if hand calculations are necessary, they are straight-

forward. Parametric multiple comparisons tests (e.g. Dunnett's) are widely available in computer software; nonparametric multiple comparisons are not, but again hand computations are straight-forward. The recommended nonparametric test is the Jonckherre-Terpstra test, followed by multiple comparisons using the nonparametric analog of Scheffe's method.

4) Choice of transformation: ANOVA is a parametric procedure based on the assumptions that observations within treatments are independent and normally distributed, and the variance of the data are homogeneous. These assumptions must be checked prior to using this approach to determine if they have been met (tests for validating these assumptions are discussed in I.3.c). When the assumptions of normality and/or homogeneity of variance are not met, transformations of the data may remedy the problem, so that the data can be analyzed by parametric, rather than nonparametric procedures. Often, but not always, the same transformation will do both. When the data include replicates of the same treatment, a Box-Cox plot can be used to choose between specific transformations to equalize variances (Box *et al.* 1978). The mean and within-group standard deviation are calculated for each group of observations. Then, y=log(SD) is plotted against x=log(mean). The slope of the linear regression indicates the transformation that best controls the variance-mean relationship. A slope near 0 indicates no transformation is required (no variance-mean relationship). A slope near 0.5 indicates a square root

transformation is needed, while a slope near 1 indicates that a log transformation is most appropriate. Often, the estimated slope will be near 1, suggesting the use of a log(x) transformation. On occasion, this may need to be modified slightly because fecundity may be zero and log(0) is undefined. The recommended transformation for the mean post-exposure fecundity is log(x+1) where x is the mean post-exposure fecundity. The recommended transformation for the ratio is log((x+1)/(y+1)), where x is the mean post-exposure fecundity and y is the mean pre-exposure fecundity.

The choice of transformation is not an issue if a nonparametric approach is chosen. The suggested nonparametric tests are based on the ranks of the observations. All the transformations considered here are monotonic; the ranks of the transformed variables are identical to the ranks of the original variables. Hence, the nonparametric analysis of transformed variables is identical to the nonparametric analysis of the untransformed variables.

c. Analysis of Variance. If a parametric analysis is utilized, one-way ANOVA can be used to test the hypothesis of no differences between treatment means. If the response is chosen to be the mean post-exposure fecundity, it should be calculated for each chamber, possibly adjusting for mortality if any occurred. The post-exposure mean fecundity in test chamber i, $\bar{x}_i$, should be transformed to $\log(\bar{x}_i + 1)$. If the

response is chosen to be the ratio, the transformed ratio should be calculated for each chamber as $\log\left(\frac{\bar{x}_i+1}{\bar{y}_i+1}\right)$, where $\bar{y}_i$ is the mean pre-exposure fecundity in chamber *i*.

For either response, treatment means and the F-test of equal means are calculated using standard computer software for one-way ANOVA. Following the ANOVA, comparisons between specific treatments can be made by fitting a concentration (dose)-response curve (e.g., a linear contrast), by testing each concentration against the control (using Dunnett's Multiple Comparisons Procedure), or comparing all pairs of treatments (using Tukey's Multiple Comparison Procedure). ANOVA is discussed in further detail in Section I.3.

d. Jonckherre-Terpstra test of ordered alternatives. The Jonckherre-Terpstra test is a nonparametric test of the hypothesis that the mean fecundities in each treatment follow a specified order. When there are three treatments (e.g., "Control," "Low," and "High" concentrations), and three treatment means (μ_C, μ_L, and μ_H), the hypothesis being tested is that $\mu_C \geq \mu_L \geq \mu_H$, with at least one strict inequality, against the null hypothesis that $\mu_C = \mu_L = \mu_H$. When the treatments are a sequence of increasing concentrations, with the concentration-response curve expected to be monotonic, a test of ordered alternatives is more powerful than the usual test of equal means (e.g., a one-way ANOVA or the nonparametric Kruskal-Wallis test).

The Jonckherre-Terpstra test can be calculated by SAS PROC FREQ (SAS Institute, Cary, NC), which can also calculate the exact p-value, the SPSS EXACT module (SPSS, Inc., Chicago, IL), and perhaps other statistical packages. It is not a test that is widely available in all statistical software. If a statistical package can calculate the Mann-Whitney (or equivalently the Wilcoxon's rank sum test) for a pair of treatments, the Jonckherre-Terpstra test statistic can be obtained from the results of Mann-Whitney tests on all pairs of treatments. If no nonparametric test is available, the Jonckherre-Terpstra test statistic can be calculated by hand.

Hand calculation of the Jonckherre-Terpstra is described and illustrated in some nonparametric statistics books (e.g., Daniel 1978). The statistic is calculated by considering all unique pairs of treatments one pair at a time. If there are three treatments, there are three pairs of treatments. Label each pair so that the higher concentration (with the lower expected fecundity) is treatment X and the lower concentration (with the higher expected fecundity) is treatment Y. Consider all possible ways that an observation in treatment X can be paired with an observation in treatment Y. If there are four observations in each treatment, there are sixteen possible pairings of observations. Count the number of pairs where the observation from treatment X is smaller than the observation from treatment Y. If two values are tied (i.e., the observations in treatment X and Y have the same fecundity), add one half for each tied pair to the count. Add these counts up over all pairs of treatments.

This sum is *J*, the test statistic for the Jonckherre-Terpstra test. A large value indicates that the fecundity tends to be larger in treatments with lower concentrations.

Exact critical values of the Jonckherre-Terpstra test statistic for a variety of sample sizes are computed in Odeh (1971) and tabulated in Daniel (1978). Critical values for the sample sizes and number of treatments used or potentially used in the test are given in Table 3. The test rejects the null hypothesis if the observed test statistic is larger than or equal to the tabulated value. The sample sizes for these tests are sufficiently small that the asymptotic Chi-square distribution should not be used.

Value of J	n=4, k=3	n=3, k=4	n=4, k=4
33	0.110		
34	0.084		
35	0.063		
36	0.046	0.117	
37	0.033	0.091	
38	0.023	0.069	
39	0.015	0.052	
40	0.0099	0.037	
41	0.0062	0.027	
42	0.0037	0.018	
43		0.012	
44		0.0080	
45		0.0050	
62			0.1058
63			0.0895
66			0.0514
67			0.0420
69			0.0272
70			0.0215
72			0.0123
73			0.0100
75			0.0056
76			0.0041

Table 3. P-values corresponding to specified values of J, the test statistic for the Jonckherre-Terpstra test of ordered alternatives (Daniel 1978). Values are provided for n=4 replicates of k=3 treatments, n=3 replicates of k=4 treatments, and n=4 replicates of k=4 treatments. Values of J for n=4, k=4 are included only when the p-value brackets a commonly used α level (e.g., the two values J=66 and J=67 bracket the 0.05 p-value).

e. Kruskal-Wallis test of equal treatment medians. An alternative to the Jonckherre-Terpstra approach is the Kruskal-Wallis test, the nonparametric equivalent of the one-way ANOVA. This test is generally available in statistical software packages. Because the sample sizes are small, exact p-values should be used instead of the asymptotic Chi-square distribution. Tables of exact p-values are available in many nonparametric statistics texts and some collections of statistical tables. Exact p-values can also be computed using SAS or SPSS (Exact module). Critical values for two common choices of α level are given in Table 4.

f. Nonparametric multiple comparisons between treatments. Nonparametric multiple comparisons between means are based on the mean ranks from the Kruskal-Wallis test. These values are usually reported in the output from Kruskal-Wallis analyses. The recommended procedure uses Scheffe's approach, adapted for the analysis of ranked data. This is appropriate for any set of comparisons between groups, including comparisons of each non-zero group to the control. For each comparison, calculate the absolute value of the difference in mean ranks, $T=|\overline{R}_i-\overline{R}_j|$. The difference between these two groups is significant if T exceeds the critical value given by:

$$\sqrt{h_{\alpha,k-1}\frac{N(N+1)}{12}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)},$$

which reduces to

$$\sqrt{h_{\alpha,k-1}\frac{k(N+1)}{6}}$$

when the two groups have the same sample size. The term $h_{\alpha,k-1}$ is the critical value for the α level Kruskal-Wallis test with k treatments where k is the number of groups and n_i is the number of observations in group i. When the sample sizes are equal, n is the number of observations per group. N is the total number of observations in the entire experiment. A subset of the critical values for multiple comparisons is given in Table 4.

		Kruskal-Wallis test		Multiple comparison	
n	k	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
4	3	5.692	7.654	6.082	7.052
3	4	7.000	8.538	7.791	8.603
4	4	7.235	9.287	9.055	10.26

Table 4. Critical values for the Kruskal-Wallis test (Daniel 1978) and Scheffe's nonparametric multiple comparisons approach with experiment-wise error rates of 5% and 1%. Values are provided for n=4 replicates of k=3 treatments, n=3 replicates of k=4 treatments, and n=4 replicates of k=4 treatments.

3. Analysis of Other Endpoints

a. Data preparation. Data on other endpoints do not provide the complexities of the fecundity or fertility data; they are measured only once at the end of the experiment (adult survival, GSI, and plasma vitellogenin and sex steroid concentrations). The data should be plotted, both as a preliminary step to help detect anomalies and unsuspected trends or patterns in the responses, and as an aid in the interpretation of the results. These plots can also be used to visually check the validity of some assumptions associated with several statistical analyses that may be used. Scatter plots of two or more variables may demonstrate the relationships among the variables, so that correlations can be observed and interactions studied.

The experimental unit in these tests is the test chamber; however, most of the endpoints are measured using individual fish. Fish-specific observations must be reduced to chamber means before use in a statistical test of treatment differences.

b. Outliers. Outliers are inconsistent or questionable data points that appear either high or low compared to the general trend exhibited by the majority of the data. Outliers can be detected by tabulating the data, plotting, or an analysis of the residuals. Explanations should be sought for any questionable data points. If the outlier cannot be linked to an explanation such as faulty equipment or human error, the outlier could be the result of a real effect. Without an explanation, data points should be discarded

only with extreme caution. If there is not an explanation, the analysis should be performed both with and without the outlier, and the results of both analyses reported. Outliers increase the variability within a sample and, therefore, decrease sensitivity of the statistical tests used.

c. Choice of analysis. The process of selecting appropriate statistical tests should follow a decision tree similar to the one illustrated in Figure 7. This approach uses tests of normality and tests of equal variance, and experimental characteristics to guide the selection of the appropriate analysis. The choice of test is based on the characteristics of each specific data set. Dunnett's Procedure and the *t*-test with Bonferroni's adjustment are parametric procedures based on the assumptions that the observations within treatments are independent and normally distributed, and that the variance of the observations is homogeneous across treatments. These assumptions should be checked prior to using the tests to determine if they have been met. If the tests fail, a nonparametric procedure such as Steel's Many-one Rank Test or Wilcoxon's Rank Sum Test may be more appropriate.

After examining the plots and descriptive statistics, assumptions of normality and homogeneity of variances among groups are tested. Normality can be tested using Shapiro-Wilks Test, among others. In general, if the data fail the test for normality, a transformation such as to log values may normalize the data. Homogeneity of variances across groups can be tested using Bartlett's Test, among others. In using

this test, it is assumed that the data are normally distributed. If data display or tests demonstrate that variance is not homogeneous across treatments, then variance-stabilizing transformations of the data may be necessary. The arcsin, square root, and logarithmic transformations are often used on dichotomous, count, and continuous data, respectively.

The choice of a particular analysis and the ability to detect departures from the assumptions of the analysis, such as the normality of the data and homogeneity of variance, is dependent on the number of replicates. One concern is that if the sample sizes are too small, the tests of equal variances and normality may have low power to detect unequal variances or non-normality, unless the deviations are very large. Because the sample sizes are relatively small under the current recommended experimental design, the data from a single study may provide relatively little information to test normality or equal variances. The statistical consequence is that tests of normality and unequal variances may have low statistical power.

If the assumptions are met, the data can be subjected to ANOVA followed by Dunnett's Multiple Comparison Procedure for comparing each of the treatment means with the control mean to determine if any of the concentrations differ from the control. It is based on the assumptions that the observations are independent and normally distributed and that the variance of the observations is homogeneous across all treatments. Dunnett's Procedure can only be used when the same number of

replicates (test chambers) have been used at each treatment. In cases where the number of replicates for different treatment levels are not equal, a *t*-test with Bonferroni's adjustment can be used as an alternative to Dunnett's Procedure. The *t*-test with Bonferroni's adjustment is based on the same assumptions of normality and homogeneity of variance as Dunnett's Procedure.

If, after suitable transformations have been carried out, the assumptions of ANOVA have not been met, nonparametric techniques should be used. Steel's Many-one Rank Test is a nonparametric test for comparing treatments with a control and is an alternative to Dunnett's Procedure. If the number of replicates (test chambers) are not equal, Wilcoxon's Rank Sum Test with Bonferroni's adjustment should be used.

If both solvent and dilution water (non-solvent) controls are included in the test, they should be compared using a *t*-test for count and continuous data and Fisher's Exact Test for categorical data. If the difference between controls is not statistically significant, then all control data can be combined for the remaining analyses.

Various software packages are available to perform these analyses such as SAS (SAS Institute, Cary, NC), SigmaStat or SYSTAT® (SPSS, Inc., Chicago, IL), ToxCalc (TidePool Scientific Software, McKinleyville, CA), or TOXSTAT® (Western Ecosystems Technology, Inc., Cheyenne, WY).

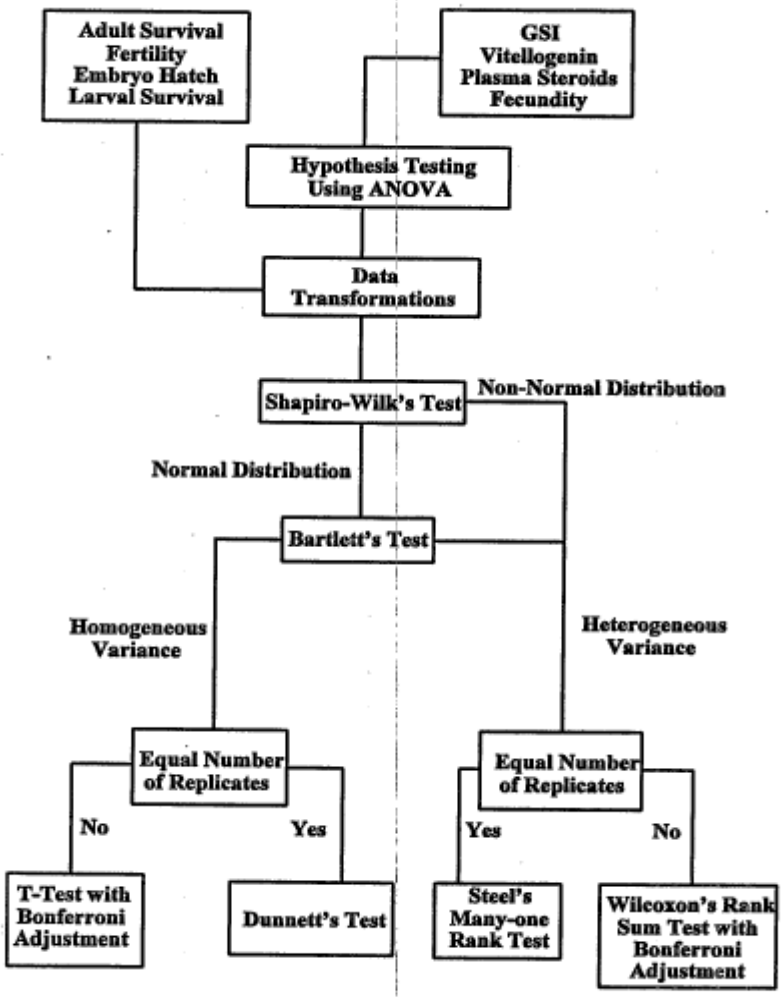

Figure 7. Flow chart for statistical analysis of data from the 21-d reproduction test with the fathead minnow.

4. Summarization

Following completion of the statistical analyses and interpretation of results, the mean and standard deviation for each endpoint should be incorporated into a summary table or figure. Statistical significance, as well as the statistical tests used, level of significance, and sample size should be indicated.

J. Test Report

A test report form must be completed for each chemical evaluated. An example of a report form for the 21-d reproduction test is provided in Appendix A (Form A-4).

1. Test Chemical

The report must include a detailed description of the test substance, including information on its CAS number, source, lot number, and purity. Additional information should be provided, when available, such as its solubility in water, K_{ow}, vapor pressure, etc. Toxicity to the fathead minnow (or other fish species) should also be reported along with pertinent information from the range-finding test.

2. Test Animals

Information must be provided on the fathead minnows used in the test. This information must include the source of the fish, age, and condition of the fish at the initiation of the test, the pre-exposure reproductive performance, the acceptability of the most recently performed reference toxicity test, and a summary of the results with the most recent performance standard test. Any observed abnormalities in reproductive behavior or performance of control fish also must be reported.

3. Test Conditions

The report must specify the conditions under which the test was performed. This must include information on the source, treatment of, and basic chemical characteristics of the dilution water. It also must include mean (± SD) and range for water temperature, dissolved oxygen, pH, hardness, alkalinity, total organic carbon, and un-ionized ammonia. The photoperiod and light intensity used during the exposure must be specified. The chamber size, water volume, flow rate, and number and composition of spawning substrates must be included. The general experimental design must be described including the number of treatments, number of replicates (test chambers) per treatment, and the number of males and females per test chamber. The report must include information on the feeding regime, including priority contaminant levels, food supplier and lot number. The basic nature of the exposure (i.e., flow-through, i.p. injection, dietary) must be stated, in addition to specific

information related to the exposure type. For example, if the exposure was a flow-through water delivery type, the daily number of volume exchanges of dilution water must be stated. If a solvent or dispersant was used to deliver the chemical, the specific solvent or dispersant and the concentrations to which the fish were exposed must be specified. If the exposure route was by i.p. injection, the carrier and injection volume must be stated. If the exposure route was dietary, the food items used to introduce the chemical must be stated in the report. The methods used for, and results (with standard deviations) of the test chemical analyses should be reported, including validation studies, and method detection/quantitation limits.

4. Results

The summary report must include the mean and standard deviation for each test endpoint. Qualitative data collected on reproductive behavior and gonadal histopathology endpoints must be summarized and reported. The results must include data for the control (plus solvent control or sham i.p.-injected control, when used) and the treatment fish. Statistical significance, as well as the statistical tests used, level of significance, and sample size must be indicated.

K. Quality Assurance

Each laboratory should incorporate a strong quality assurance (QA) program from the start of testing. Prior to the actual initiation of the pre-exposure period, Standard Operating Procedures

(SOPs) should be prepared for all aspects of the study, including culturing of test animals, operation of the test system, generation of stock solutions, instrument calibration, analysis of test chemical, analyses of water quality characteristics, sample tracking and chain-of-custody procedures, performance of pre-exposure test, and performance of the 21-d reproduction test. Staff that will be performing tasks on the project should be familiar with the SOPs before the tests begin.

QA practices must address all activities that affect the quality of the final data obtained from the test, such as: (1) test substance, (2) source and condition of the test animals, (3) water quality characteristics, (4) condition of exposure apparatus and equipment, (5) maintenance of chemical exposure concentrations, (6) analytical methodology, (7) instrument calibration, (8) replication, (9) record keeping, and (10) data evaluation and interpretation. During routine activities, established quality control (QC) practices should be followed which ensure generation of data that are of known quality.

The fish used in the test should appear healthy, behave normally, feed well, reproduce successfully, and have low mortality in culture, pre-exposure period, and in the test controls. Routine water quality characteristics such as temperature, dissolved oxygen, pH, hardness, alkalinity, total organic carbon, and un-ionized ammonia should be monitored to assure quality of the system used to maintain the animals. All instruments should be calibrated and standardized following the instrument manufacturer's procedures, and monitored according to standard methods.

Proper record keeping is important. A complete file should be maintained for each test indicating test chemical, investigator, and dates of initiation and termination of test. The file should contain the original data sheets from the test; source and information about the test chemical and concentration of any solvent used; detailed records of the animals used in the test, such as source, age, and any other pertinent information; test conditions employed; description of the experimental design; methods used for, and results of, analysis of the test chemical; source and characteristics of dilution water; information on the calibration of equipment and instruments; and the methods used for, and results of, statistical analyses of the data. The file should also contain anything unusual about the test, any deviations from established SOPs, and all other relevant information. Laboratory data should be recorded on a real-time basis to prevent loss of information or inadvertent introduction of errors into the record. Original data sheets should be signed and dated by the laboratory personnel performing the tests.

Internal audits should be performed by the testing laboratory's QA officer during the actual performance of each test by inspection of the SOPs, evaluation of the performance of the test system, and inspection of the biological and chemical record books. Post-test audits should be performed of the data analysis, statistical treatment, and interpretation of results, with records made of any deviations from the test protocols. QA records should address test animal health, pre-exposure reproductive performance, the performance of the test from a biological perspective, and the performance of the test from a chemical perspective. A QA summary should be prepared by the QA officer for each test and should accompany the test report.

L. Interpretation of Results

The specific test described in this document, as well as relatively similar reproduction studies with the fathead minnow (e.g., Kramer *et al.* 1998; Giesy *et al.* 2000; Harries *et al.* 2000) have been conducted with a number of EDCs representative of several MOA. In this section we provide an overview of responses in the test to chemicals with known endocrine MOA, as a basis for interpretation of study results with unknown chemicals. Some individual responses are very diagnostic in terms of identification of a specific endocrine MOA (e.g., induction of vitellogenin in males caused by estrogen receptor agonists), but in many cases it is/will prove necessary to consider patterns of responses in the whole suite of endpoints to assess which (if any) endocrine pathway has been affected. It must be noted that the database from which this interpretive guidance was developed is limited. For example, tests with chemicals with mixed (endocrine) MOA have been rare, and likely would result in unanticipated patterns of responses (e.g., see methyltestosterone example below). Another important shortcoming in the current knowledge base is a lack of data for chemicals which affect reproduction, but not through alterations in the endocrine systems of concern. The assumption in these cases is that some generic measure of reproductive potential would be affected (e.g., fecundity, GSI) in the absence of changes in other, more diagnostic, endpoints such as secondary sex characteristics, plasma vitellogenin and sex steroid concentrations, and gonadal histopathology.

Table 5 summarizes responses of fathead minnows to different EDCs in the context of the suite of endpoints described in this document. The most work, by far, has been with estrogen receptor

agonists. Strong agonists, such as β-estradiol, reduce fecundity of actively-spawning animals, and consistently induce vitellogenin in males (Table 5; Kramer *et al.* 1998; Panter *et al.* 1998; 2002; Tyler *et al.* 1999; Korte *et al.* 2000). Other endpoints that have been reported to be affected by strong estrogen receptor agonists in sexually-mature fathead minnows include gonadal (testicular and ovarian) histopathology and alterations in secondary sex characteristics (Panter *et al.* 1998; Miles-Richardson *et al.* 1999a; Harries *et al.* 2000). Exposure of fathead minnows to chemicals that are weaker estrogen receptor agonists (e.g., alkylphenols, methoxychlor) elicit a qualitatively similar pattern of effects similar to those observed after exposure to stronger agonists, although the magnitude of the effects (not surprisingly) differs between weak and strong estrogens (Miles-Richardson *et al.* 1999b; Giesy *et al.* 2000; Harries *et al.* 2000; Ankley *et al.* 2001). For example, methoxychlor significantly decreased (but did not completely inhibit) spawning of fathead minnows at a concentration of about 5 µg/L (Ankley *et al.* 2001). At this concentration, a significant induction of vitellogenin in male fathead minnows was observed; however, the response was much less pronounced than when adult male fathead minnows were exposed to strongly estrogenic substances (Panter *et al.* 1998; Korte *et al.* 2000; Ankley *et al.* 2001). There also have been descriptions of alterations in secondary sex characteristics and ovarian histopathology in adult fathead minnows exposed to weak estrogens (Miles-Richardson *et al.* 1999b; Harries *et al.* 2000; Ankley *et al.* 2001). Plasma concentrations of sex steroids also can be affected (in a sex-specific manner) by weak estrogen receptor agonists (Table 5; Ankley *et al.* 2001); presumably, if comparable data were available, this also would be observed in exposures with strong estrogens.

One androgen receptor agonist, the synthetic compound methyltestosterone, has been evaluated using the short-term fathead minnow reproduction test (Table 5; Ankley *et al.* 2001). It appeared that the methyltestosterone elicited a suite of responses indicative of a chemical with a mixed estrogenic and androgenic MOA rather than a "pure" androgen. Exposure to methyltestosterone at concentrations $\geq$ 0.2 mg/L caused an immediate cessation of spawning. Consistent with previous demonstrations (from aquaculture studies) that methyltestosterone is androgenic in fish, the adult females were clearly masculinized, exhibiting pronounced nuptial tubercle development within about 6 d of exposure. However, methyltestosterone also caused a large induction of vitellogenin in both males and females, which is a response consistent with (and relatively specific to) an estrogen receptor agonist. This may have occurred because methyltestosterone can be converted via aromatase to a methyl-estradiol analogue (Dr. D. Kime, University of Sheffield, personal communication), which would have resulted in the fish actually being exposed to an estrogen/androgen mixture. Given this, it is difficult to say whether other responses observed in the test (e.g., reduced steroid concentrations, reduced GSI, abnormal gonadal histology; Table 5) were due to the androgenic or estrogenic (or combined) nature of methyltestosterone. Based, however, upon this study and early work by Smith (1974) with fathead minnows exposed to known androgens, masculinization of adult females would appear to be a very diagnostic response for this MOA.

Two putative anti-androgens have been evaluated in 21-d reproduction studies with the fathead minnow (Table 5). Makynen *et al.* (2000) assessed the effects of vinclozolin on fathead minnows held in a paired-breeding situation. Due to minimal reproduction in controls from that

experiment, it was difficult to determine whether exposure to the mammalian androgen receptor antagonist affected fecundity of the fish. Vinclozolin did not markedly affect plasma steroid concentrations in males or females, and vitellogenin was not measured in that experiment. At a concentration of about 700 µg/L vinclozolin did, however, cause a significant reduction in GSI of the females, which was accompanied by retarded oocyte maturation and atresia. Because neither vinclozolin or its primary metabolites bound to the fathead minnow androgen receptor *in vitro* (Makynen *et al.* 2000), it was uncertain whether the responses observed in the gonads of the females were truly indicative of an anti-androgen. Therefore, the results of a reproduction test with the androgen receptor antagonist flutamide (which does bind to the fathead minnow androgen receptor; Makynen *et al.* 2000) may be more descriptive of the expected pattern of responses associated with exposure of reproductively-active fathead minnows to an anti-androgen. At concentrations ranging from 60 to 600 µg/L, flutamide caused a concentration-dependent decrease in fecundity which, as was observed with vinclozolin, was accompanied by decreased GSI and increased oocyte atresia/retarded maturation in the female fathead minnow (Table 5; Jensen *et al.* 2002). Flutamide also affected steroid concentrations in both sexes and appeared to cause a slight increase in vitellogenin concentrations in the female. In addition, flutamide exposure resulted in subtle indications of gonadal histopathology in the males, which was comprised of germ cell necrosis and reduced spermatogenesis. Based on these results, the most consistent effects of the anti-androgens in this test appear to be expressed in the gonads of the females (GSI, histopathology).

Aromatase (CYP19) is a cytochrome P450-based enzyme that, under normal physiological conditions, converts testosterone to β-estradiol. There is emerging evidence that the MOA via which some EDCs exert their effects is through alterations in steroid synthesis associated with inhibition of aromatase activity (i.e., CYP19). Fadrozole, a classical inhibitor of aromatase activity, was evaluated using the protocol described in this document (Table 5; Ankley *et al.* 2002). The chemical caused a concentration-dependent reduction in fecundity at concentrations ranging from about 1.5 to 50 µg/L. Consistent with the presumed MOA, there also was a concentration-dependent decrease in both plasma β-estradiol and vitellogenin in the female fathead minnow. In addition, plasma concentrations of testosterone and 11-ketotestosterone were increased in the males, and histological alterations observed in the gonads of both sexes. Given the specificity of aromatase inhibitors, the decreases in β-estradiol and, subsequently, vitellogenin in the female fathead minnow should be an excellent diagnostic response for this class of EDCs. Previous studies with fish have emphasized vitellogenin induction in males as a highly-specific indicator of an endocrine MOA (estrogen receptor agonists); these data indicate an equally useful and diagnostic response associated with vitellogenin reductions in (sexually-mature) females. This endpoint presumably would reflect effects of chemicals, not only on β-estradiol synthesis (as for fadrozole), but the action of chemicals that act as estrogen receptor antagonists (Panter *et al.* 2002)

The patterns of responses summarized in Table 5 clearly represent only a small subset of possible outcomes associated with exposure to EDCs. Adverse effects associated with some MOAs (estrogen receptor agonists, androgen receptor agonists, and aromatase inhibitors) should be

easily identified. Identification of chemicals as anti-androgens may be more equivocal (although these types of chemicals would clearly be "flagged" as endocrine-active through alterations in gonadal histology and, perhaps, steroid concentrations). As this test is conducted with additional chemicals reflective of the MOA discussed above, as well as other MOA, guidance in interpreting test results will expand.

Table 5. Performance of fathead minnow screening tests with known endocrine-disrupting chemicals.

MOA/Chemical	Effects Observed						
	Fecundity	Sex	Secondary Sex Characteristics	Plasma Steroids	Vitellogenin	GSI	Gonadal Histology
ER Agonist							
β-Estradiol [a]	↓	*Male*	↓ Tubercle size	ND [b]	↑	↓	Sertoli cell proliferation, degenerative changes
		Female	None	ND	↑	ND	Oocyte atresia, retarded oocyte maturation
Methoxychlor [c]	↓	*Male*	None	↓ T, KT	↑	None	None
		Female	None	↓ E	None	None	Oocyte atresia
AR Agonist							
Methyltestosterone [c]	↓	*Male*	↑ Tubercle size	↓ T, KT	↑	↓	Germ cell necrosis, reduced spermatogenesis
		Female	Masculinized	↓ E, T	↑	↓	Oocyte atresia, retarded oocyte maturation
AR Antagonist							
Vinclozolin [d]	ND	*Male*	None	↑ E	ND	None	None
		Female	None	None	ND	↓	Retarded oocyte maturation
Flutamide [e]	↓	*Male*	None	↑ KT	None	None	Germ cell necrosis, reduced spermatogenesis
		Female	None	↓ E ↑ T	↑	↓	Oocyte atresia, retarded oocyte maturation
Aromatase Inhibitor							
Fadrozole [f]	↓	*Male*	None	↑ T, KT	None	↑	Increased spermatogenesis
		Female	None	↓ E	↓	None	Oocyte atresia, retarded oocyte maturation

[a] Data obtained from Kramer *et al.* 1998; Panter *et al.* 1998; 2002; Miles-Richardson *et al.* 1999a; and Korte *et al.* 2000.
[b] Not determined.
[c] Data obtained from Ankley *et al.* 2001.
[d] Data obtained from Makynen *et al.* 2000.
[e] Data obtained from Jensen *et al.* 2002.
[f] Data obtained from Ankley *et al.* 2002.

M. References

American Society for Testing and Materials. 2000a. Standard guide for conducting acute toxicity tests with fishes, macroinvertebrates, and amphibians. Method E 729-96. In *Annual Book of ASTM Standards*. Volume 11.05. West Conshohocken, PA, USA, pp 213-233.

American Society for Testing and Materials. 2000b. Standard guide for conducting early life-stage toxicity tests with fishes. Method E 1241-98. In *Annual Book of ASTM Standards*. Volume 11.05. West Conshohocken, PA, USA, pp 941-967.

Andreassen TK, Korsgaard B. 2000. Characterization of a cytosolic estrogen receptor and its up-regulation by 17β-estradiol and the xenoestrogen 4-*tert*-octylphenol in the liver of the eelpout (*Zoarces viviparus*). *Comp Biochem Phys* C 125:299-313.

Ankley GT, Giesy JP. 1998. Endocrine disruptors in wildlife: A weight of evidence perspective. In Kendall R, Dickerson R, Suk W, Giesy J, eds, *Principles and Processes for Assessing Endocrine Disruption in Wildlife*. SETAC, Pensacola, FL, USA, pp 349-368.

Ankley GT, Johnson RD, Detenbeck NE, Bradbury SP, Folmar L, Toth G. 1997. Development of a research strategy for assessing the ecological risk of endocrine disruptors. *Rev Toxicol* 1:71-106.

Ankley GT, *et al.* 1998a. Overview of a workshop on screening methods for detecting potential (anti-) estrogenic/androgenic chemicals in wildlife. *Environ Toxicol Chem* 17:68-87.

Ankley GT, Tietge JE, DeFoe DL, Jensen KM, Holcombe GW, Durhan EJ, Diamond SA. 1998b. Effects of ultraviolet light and methoprene on survival and development of *Rana pipiens*. *Environ Toxicol Chem* 17:2530-2542.

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.

Ankley GT, Kahl MD, Jensen KM, Hornung MW, Korte JJ, Makynen EA, Leino RL. 2002. Evaluation of the aromatase inhibitor fadrozole in a short-term reproduction assay with the fathead minnow (*Pimephales promelas*). *Toxicol Sci* 67:121-130.

Bardach JE, Bernstein JJ, Hart JS, Brett JR. 1966. Tolerance to temperature extremes: Animals. Part IV. Fishes. In Altman PL, Dittmer D, eds, *Environmental Biology*. Fed. Am. Soc. Exp. Biol., Bethesda, MD, USA, pp 37-80.

Benoit DA, Mattson VR, Olson DL. 1982. A continuous-flow mini-diluter system for toxicity testing. *Water Res* 16:457-464.

Box GEP, Hunter WG, Hunter JS. 1978. *Statistics for Experimenters*. John Wiley, New York, NY, USA, 653 pp.

Bradbury SP. 1994. Predicting modes of toxic action from chemical structure: An overview. *SAR QSAR Environ Res* 2:89-104.

Bradbury SP. 1995. Quantitative structure-activity relationships and ecological risk assessment: An overview of predictive aquatic toxicology research. *Toxicol Lett* 79:229-237.

Bradbury SP, Lipnick RL. 1990. Introduction: Structural properties for determining mechanisms of toxic action. *Environ Health Perspect* 87:181-182.

Brooke LT, Call DJ, Geiger DL, Northcott CE, eds. 1984. *Acute Toxicities of Organic Chemicals to Fathead Minnows (Pimephales promelas)*, Vol I. Center for Lake Superior Environmental Studies, University of Wisconsin, Superior, WI, USA, 414 pp.

Brungs WA. 1971. Chronic effects of constant elevated temperature on the fathead minnow (*Pimephales promelas* Rafinesque). *Trans Am Fish Soc* 100:659-664.

Call DJ, Geiger DL, eds. 1992. *Subchronic Toxicities of Industrial and Agricultural Chemicals to Fathead Minnows (Pimephales promelas)*, Vol I. Center for Lake Superior Environmental Studies, University of Wisconsin, Superior, WI, USA.

Call DJ, Brooke LT, Knuth ML, Poirier SH, Hoglund MD. 1985. Fish subchronic toxicity prediction model for industrial organic chemicals that produce narcosis. *Environ Toxicol Chem* 4:335-341.

Call DJ, Markee TP, Geiger DL, Brooke LT, VandeVenter FA, Cox DA, Genisot KI, Robillard KA, Gorsuch JW, Parkerton TF, Reiley MC, Ankley GT, Mount DR. 2001. An assessment of the toxicity of phthalate esters to freshwater benthos. I. Aqueous exposures. *Environ Toxicol Chem* 20:1798-1804.

Chang CF, Hung CY, Chiang MC, Lan SC. 1999. The concentrations of plasma sex steroids and gonadal aromatase during controlled sex differentiation in grey mullet, *Mugil cephalus*. *Aquaculture* 177:37-45.

Colborn T, Dumanoski D, Meyers JP. 1996. *Our Stolen Future: Are We Threatening Our Fertility, Intelligence, and Survival? A Scientific Detective Story*. Dutton, New York, NY, USA.

Crim LW, Glebe BD. 1990. Reproduction. In Schreck CB, Moyle PB, eds, *Methods for Fish Biology*. American Fisheries Society, Bethesda, MD, USA, pp 529-553.

Crisp TM, *et al.* 1997. Special report on environmental endocrine disruption: An effects assessment and analysis. EPA/630/R-96/012. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC, USA.

Daniel WW. 1978. *Applied Nonparamentric Statistics.* Houghton Mifflin Co., Boston, MA, USA. 503 pp.

DeFoe DL. 1975. Multichannel toxicant injection system for flow-through bioassays. *J Fish Res Board Can* 32:544-546.

DeFoe D, Holcombe G. 1997. An innovative computer-driven, electronic diluter for flow-through toxicity tests. *Abstracts*, 18th Annual Meeting, Society of Environmental Toxicology and Chemistry, San Francisco, CA, USA, p 305.

Devine G. 1968. A study of the smallmouth bass in ponds with special consideration of minnows and decapods as forage. MA Thesis, University of Missouri, Columbia, MO, USA, 115 pp.

DeVito M, *et al.* 1999. Screening methods for thyroid hormone disruptors. *Environ Health Perspect* 107:407-415.

Elonen GE, Spehar, RL, Holcombe GW, Johnson RD, Fernandez JD, Erickson RJ, Tietge JE, Cook PM. 1998. Comparative toxicity of 2,3,7,8-tetrachloro-*p*-dioxin to seven freshwater fish species during early life-stage development. *Environ Toxicol Chem* 17:472-483.

Erickson RJ, Ankley GT, DeFoe DL, Kosian PA, Makynen EA. 1999. Additive toxicity of binary mixtures of phototoxic polycyclic aromatic hydrocarbons to the oligochaete *Lumbriculus variegatus*. *Toxicol Appl Pharmacol* 154:97-105.

Fernandez JD, Denny JS, Tietge JE. 1998. A simple apparatus for administering 2,3,7,8-tetrachlorodibenzo-*p*-dioxin to commercially available pelletized fish food. *Environ Toxicol Chem* 17:2058-2062.

Flickinger SA. 1969. Determination of sexes in the fathead minnow. *Trans Amer Fish Soc* 98:526-527.

Gale WF, Buynak GL. 1982. Fecundity and spawning frequency of the fathead minnow - A fractional spawner. *Trans Am Fish Soc* 111:35-40.

Gast MH, Brungs WA. 1973. A procedure for separating eggs of the fathead minnow. *Prog Fish-Cult* 35:54.

Geiger DL, Northcott CE, Call DJ, Brooke LT, eds. 1985. *Acute Toxicities of Organic Chemicals to Fathead Minnows (Pimephales promelas)*, Vol II. Center for Lake Superior Environmental Studies, University of Wisconsin, Superior, WI, USA, 326 pp.

Geiger DL, Poirier SH, Brooke LT, Call DJ, eds. 1986. *Acute Toxicities of Organic Chemicals to Fathead Minnows (Pimephales promelas)*, Vol III. Center for Lake Superior Environmental Studies, University of Wisconsin, Superior, WI, USA, 328 pp.

Geiger DL, Call DJ, Brooke LT, eds. 1988. *Acute Toxicities of Organic Chemicals to Fathead Minnows (Pimephales promelas)*, Vol IV. Center for Lake Superior Environmental Studies, University of Wisconsin, Superior, WI, USA, 355 pp.

Geiger DL, Brooke LT, Call DJ, eds. 1990. *Acute Toxicities of Organic Chemicals to Fathead Minnows (Pimephales promelas)*, Vol V. Center for Lake Superior Environmental Studies, University of Wisconsin, Superior, WI, USA, 332 pp.

Gelinas D, Pitoc GA, Callard GV. 1998. Isolation of a goldfish brain cytochrome P450 aromatase cDNA: mRNA expression during the seasonal cycle and after steroid treatment. *Mol Cell Endocrinol* 138:81-93.

Giesy JP, Pierens SL, Snyder, EM, Miles-Richardson S, Kramer VJ, Snyder SA, Nichols KM, Villeneuve DA. 2000. Effects of 4-nonylphenol on fecundity and biomarkers of estrogenicity in fathead minnows (*Pimephales promelas*). *Environ Toxicol Chem* 19:1368-1377.

Gray LE, *et al.* 1997. Endocrine screening methods workshop report: Detection of estrogenic and androgenic hormonal and antihormonal activity for chemicals that act via receptor and steroidogenic enzyme mechanisms. *Reprod Toxicol* 11:719-750.

Harries JE, Runnalls T, Hill E, Harris CA, Maddix S, Sumpter JP, Tyler CR. 2000. Development of a reproductive performance test for endocrine disrupting chemicals using pair-breeding fathead minnows (*Pimephales promelas*). *Environ Sci Technol* 34:3003-3011.

Held JW, Peterka JJ. 1974. Age, growth and food habits of the fathead minnow, *Pimephales promelas*, in North Dakota saline lakes. *Trans Am Fish Soc* 103:743-756.

Hoke RA, Ankley GT, Cotter AM, Goldenstein T, Kosian PA, Phipps GL, Vandermeiden FM. 1994. Evaluation of equilibrium partitioning theory for predicting acute toxicity of field-collected sediments contaminated with DDT, DDE and DDD to the amphipod *Hyalella azteca*. *Environ Toxicol Chem* 13:157-166.

Hoke RA, Kosian PA, Ankley GT, Cotter AM, Vandermeiden FM, Phipps GL, Durhan EJ. 1995. Check studies with *Hyalella azteca* and *Chironomus tentans* in support of the development of a sediment quality criterion for dieldrin. *Environ Toxicol Chem* 14:435-443.

Jarvinen AW, Ankley GT. 1999. *Linkage of Effects to Tissue Residues: Development of a Comprehensive Database for Aquatic Organisms Exposed to Inorganic and Organic Chemicals.* SETAC, Pensacola, FL, USA, 364 pp.

Jensen KM, Korte JJ, Kahl MD, Pasha MS, Ankley GT. 2001. Aspects of basic reproductive biology and endocrinology in the fathead minnow (*Pimephales promelas*). *Comp Biochem Physiol* C 128:127-141.

Jensen KM, Kahl MD, Korte JJ, Makynen EA, Leino RL, Ankley GT. 2002. Effects of flutamide in a short-term reproduction assay with the fathead minnow (*Pimephales promelas*). In preparation.

Kahl MD, Makynen EA, Kosian PA, Ankley GT. 1997. Toxicity of 4-nonylphenol in a life-cycle test with the midge *Chironomus tentans*. *Ecotoxicol Environ Saf* 38:155-160.

Kahl MD, Russom CL, DeFoe DL, Hammermeister DE. 1999. Saturation units for use in aquatic bioassays. *Chemosphere* 39:539-551.

Kahl MD, Jensen KM, Korte JJ, Ankley GT. 2001. Effects of handling on endocrinology and reproductive performance of the fathead minnow. *J Fish Biol* (in press).

Kavlock RJ, *et al.* 1996. Research needs for the risk assessment of health and environmental effects of endocrine disruptors: A report of the U.S. EPA-sponsored workshop. *Environ Health Perspect* 104:715-740.

Keith LH. 1997. *Environmental Endocrine Disruptors: A Handbook of Property Data.* John Wiley & Sons, New York, NY, USA.

Kenaga EE. 1982. Predictability of chronic toxicity from acute toxicity of chemicals in fish and aquatic invertebrates. *Environ Toxicol Chem* 1:347-358.

Kime DE. 1998. *Endocrine Disruption in Fish.* Kluwer Academic Publishers, Norwell, MA, USA, 396 pp.

Kitano T, Takamune K, Nagahama Y, Abe SI. 2000. Aromatase inhibitor and 17α-methyltestosterone cause sex reversal from genetical females to phenotypic males and suppression of P450 aromatase gene expression in Japanese flounder (*Paralichthys olivaceus*). *Mol Reprod Develop* 56:1-5.

Korte JJ, Kahl MD, Jensen KM, Pasha MS, Parks LG, LeBlanc GA, Ankley GT. 2000. Fathead minnow vitellogenin: complementary DNA sequence and messenger RNA and protein expression after 17β-estradiol treatment. *Environ Toxicol Chem* 19:972-981.

Kramer VJ, Miles-Richardson S, Pierens SL, Giesy JP. 1998. Reproductive impairment and induction of alkaline-labile phosphate, a biomarker of estrogen exposure, in fathead minnows (*Pimephales promelas*) exposed to waterborne 17β-estradiol. *Aquat Toxicol* 40:335-360.

Länge R, Hutchinson TH, Croudace CP, Siegmund F, Schweinfurth H, Hampe P, Panter GH, Sumpter JP. 2001. Effects of the synthetic estrogen 17α-ethinylestradiol on the life-cycle of the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 20:1216-1227.

Lewis PA, Klemm DJ, Lazorchak JM, Norberg-King TJ, Peltier WH, Heber MA. 1994. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. 3rd ed. EPA/600/4-91/002. U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Cincinnati, OH.

Lewis RJ, Sr. 1991. *Reproductively Active Chemicals. A Reference Guide.* Van Nostrand Reinhold, New York, NY, USA.

Makynen EA, Kahl MD, Jensen KM, Tietge JE, Wells KL, Van Der Kraak G, Ankley GT. 2000. Effects of the mammalian antiandrogen vinclozolin on development and reproduction of the fathead minnow (*Pimephales promelas*). *Aquat Toxicol* 48:461-475.

Mayer FL, Jr, Ellersieck MR. 1986. Manual of acute toxicity: Interpretation and data base for 410 chemicals and 66 species of freshwater animals. Resource Publication No. 160. U.S. Fish and Wildlife Service, Washington, DC, USA, 506 pp.

McCarraher DB, Thomas R. 1968. Some ecological observations on the fathead minnow, *Pimephales promelas*, in the alkaline waters of Nebraska. *Trans Am Fish Soc* 97:52-55.

McKim JM. 1977. Evaluation of tests with early life stages of fish for predicting long-term toxicity. *J Fish Res Board Can* 34:1148-1154.

Melo AC, Edmunds JSG, Ramsdell JS. 1999. Localization and sex steroid regulation of brain aromatase activity: identification of a positive feedback mechanism that may be triggered by environmental cues. In Henshel DS, Black MC, Harrass MC, eds, *Environmental Toxicology and Risk Assessment: Standardization of Biomarkers for Endocrine Disruption and Environmental Assessment: 8th Volume, ASTM STP 1364.* American Society for Testing and Materials, West Conshohocken, PA, USA, pp 226-238.

Miles-Richardson SR, Kramer VJ, Fitzgerald SD, Render JA, Yamini B, Barbee SJ, Giesy JP. 1999a. Effects of waterborne exposure of 17β-estradiol on secondary sex characteristics and gonads of fathead minnows (*Pimephales promelas*). *Aquat Toxicol* 47:129-145.

Miles-Richardson SR, Pierens SL, Nichols KM, Kramer VJ, Snyder EM, Snyder SA, Render JA, Fitzgerald SD, Giesy JP. 1999b. Effects of waterborne exposure to 4-nonylphenol and

nonylphenol ethoxylate on secondary sex characteristics and gonads of fathead minnows (*Pimephales promelas*). *Environ Res* 80:S122-S137.

Mommsen TP, Walsh PJ. 1988. Vitellogenesis and oocyte assembly. In Hoar WS, Randall VJ eds., *Fish Physiology*, Vol. XIA. Academic Press, San Diego, CA, USA, pp 347-406.

Mount DI. 1973. Chronic effect of low pH on fathead minnow survival, growth and reproduction. *Water Res* 7:987-993.

Mount DI, Brungs WA. 1967. A simplified dosing apparatus for fish toxicology studies. *Water Res* 1:21-29.

Odeh RE. 1971. On Jonckherre's k-sample test against ordered alternatives. *Technometrics* 13:912-918.

Pakdel F, Féon S, Le Gac F, Le Menn F, Valotaire Y. 1991. In vivo estrogen induction of hepatic estrogen receptor mRNA and correlation with vitellogenin mRNA in rainbow trout. *Mol Cell Endocrinol* 75:205-212.

Panter GH, Thompson RS, Sumpter JP. 1998. Adverse reproductive effects in male fathead minnows (*Pimephales promelas*) exposed to environmentally relevant concentrations of the natural oestrogens, oestradiol and oestrone. *Aquat Toxicol* 42:243-253.

Panter GH, Hutchinson TH, Länge R, Lye CM, Sumpter JP, Zerulla M, Tyler CT. 2001. Utility of a juvenile fathead minnow screening assay for detecting (anti)-estrogenic substances. *Environ Toxicol Chem* (in press).

Parks LG, Cheek AO, Denslow ND, Heppell SA, McLachlan JA, LeBlanc GA, Sullivan CV. 1999. Fathead minnow (*Pimephales promelas*) vitellogenin: purification, characterization and quantitative immunoassay for the detection of estrogenic compounds. *Comp Biochem Physiol C* 123:113-125.

Pelissero C, Lenczowski MJP, Chinzi D, Davail-Cuisset B, Sumpter JP, Fostier A. 1996. Effects of flavonoids on aromatase activity, an *in vitro* study. *J Steroid Biochem Molec Biol* 57:215-223.

Phipps GL, Holcombe GW. 1985. A method for aquatic multiple species toxicant testing: acute toxicity of 10 chemicals to 5 vertebrates and 2 invertebrates. *Environ Pollut (Ser A)* 38:141-157.

Pillard DA. 1995. Comparative toxicity of formulated glycol deicers and pure ethylene and propylene glycol to *Ceriodaphnia dubia* and *Pimephales promelas*. *Environ Toxicol Chem* 14:311-315.

Poirier SH, Knuth ML, Anderson-Buchou CD, Brooke LT, Lima AR, Shubat PJ. 1986. Comparative toxicity of methanol and N,N-dimethylformamide to freshwater fish and invertebrates. *Bull Environ Contam Toxicol* 37:615-621.

Rawson DS, Moore JE. 1944. The saline lakes of Saskatchewan. *Cancer Res* 22:141-201.

Russom CL, Anderson EB, Greenwood BE, Pilli A. 1991. ASTER: an integration of the AQUIRE data base and the QSAR system for use in ecological risk assessments. *Sci Total Environ* 109/110:667-670.

Russom CL, Bradbury SP, Broderius SJ, Hammermeister DE, Drummond RA. 1997. Predicting modes of toxic action from chemical structure: acute toxicity in the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 16:948-967.

Scholz S, Gutzeit HO. 2000. 17α-ethinylestradiol affects reproduction, sexual differentiation and aromatase gene expression of the medaka (*Oryzias latipes*). *Aquat Toxicol* 50:363-373.

Shilling AD, Carlson DB, Williams DE. 1999. Rainbow trout, *Oncorhynchus mykiss*, as a model for aromatase inhibition. *J Steroid Biochem Molec Bio* 70:89-95.

Sharp JR. 1991. The influence of toxicants on the teratogenic response of fishes: a guide to the literature, a glossary of terms and a pictorial atlas. *Trace Substances in Environmental Health* 24:63-80.

Smith, RJF. 1974. Effects of 17α-methyltestosterone on the dorsal pad and tubercles of fathead minnows (*Pimephales promelas*). *Can J Zool* 52:1031-1038.

Smith, RJF. 1978. Seasonal changes in the histology of the gonads and dorsal skin of the fathead minnow, *Pimephales promelas*. *Can J Zool* 56:2103-2109.

Society of Environmental Toxicology and Chemistry-Europe. 1997. Endocrine modulators and wildlife: assessment and testing. In Tattersfield L, Matthiessen P, Campbell PM, Grandy N, Länge R, eds. *SETAC Europe/OECD/EC Expert Workshop on Endocrine Modulators, Wildlife-Assessment and Testing*. SETAC-Europe Press, Brussels, Belgium, 126 pp.

Sprague JB. 1969. Measurement of pollutant toxicity to fish. I. Bioassay methods for acute toxicity. *Water Res* 3:793-821.

Sumpter JP, Jobling S. 1995. Vitellogenesis as a biomarker for estrogenic contamination of the aquatic environment. *Environ Health Perspect* 103:173-178.

Suter GW II, Rosen AE, Linuer E, Parkhurst DF. 1987. Endpoints for responses of fish to chronic toxic exposures. *Environ Toxicol Chem* 6:793-809.

Thompson EA Jr, Siiteri PK. 1974. Utilization of oxygen and reduced nicotinamide adenine dinucleotide phosphate by human placental microsomes during aromatization of androstenedione. *J Biol Chem* 249:5364-5372.

Todo T, Ikeuchi T, Kobayashi T, Nagahama Y. 1999. Fish androgen receptor: cDNA cloning, steroid activation of transcription in transfected mammalian cells, and tissue mRNA levels. *Biochem Biophys Res Comm.* 254:378-383.

Tyler CR, van Aerle R, Hutchinson TH, Maddix S, Trip H. 1999. An *in vivo* testing system for endocrine disruptors in fish early life stages using induction of vitellogenin. *Environ Toxicol Chem* 18:337-347.

U.S. Environmental Protection Agency. 1976. *Quality Criteria for Water* (Red Book). Office of Water Planning and Standards, Washington, DC, USA.

U.S. Environmental Protection Agency. 1982. User's guide for conducting life-cycle chronic toxicity tests with fathead minnows. EPA/600/8-81-011. U.S. Environmental Protection Agency, Environmental Research Laboratory, Duluth, MN, USA.

U.S. Environmental Protection Agency. 1987. Guidelines for the culture of fathead minnows *Pimephales promelas* for use in toxicity tests. EPA/600/3-87/001. U.S. Environmental Protection Agency, Environmental Research Laboratory, Duluth, MN, USA.

U.S. Environmental Protection Agency. 1993. *Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms*, 4th ed, Weber CI, ed. EPA-600/4-90-027F. Office of Research and Development, Environmental Monitoring Systems Laboratory, Cincinnati, OH, USA.

U.S. Environmental Protection Agency. 1994. *Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Water to Freshwater Organisms*, 3rd ed, Lewis PA, Klemm DJ, Lazorchak JM, Norberg-King TJ, Peltier WH, Heber MA, eds. EPA-600-4-91-002. Office of Research and Development, Environmental Monitoring Systems Laboratory, Cincinnati, OH, USA.

U.S. Environmental Protection Agency. 1996a. Strategic Plan for the Office of Research and Development. EPA/600/R3-91-063. Washington, DC, USA.

U.S. Environmental Protection Agency. 1996b. Prehatching development of the fathead minnow *Pimephales promelas* Rafinesque. EPA/600/R-96/079. Washington, DC, USA.

U.S. Environmental Protection Agency. 1998. Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) report. Office of Prevention, Pesticides and Toxic Substances, Washington, DC, USA.

Veith GD, Broderius SJ. 1990. Rules for distinguishing toxicants that cause Type I and Type II narcosis syndromes. *Environ Health Perspect* 87:207-211.

APPENDIX A

DATA REPORTING FORMS

Form A-1. Record of Adult Fish Survival and Water Quality Characteristics

Form A-2. Record of Spawning Activity

Form A-3. Fathead Minnow Sample Record

Form A-4. Report Form for the 21-d Reproduction Test

APPENDIX A-1

RECORD OF ADULT FISH SURVIVAL AND WATER QUALITY CHARACTERISTICS

STUDY: _____

Date: _____ **Initials:** _____

Mortality

1-A	1-B	1-C	1-D	2-A	2-B	2-C	2-D	3-A	3-B	3-C	3-D

Comments:

Water Quality Characteristics

	1-A	1-B	1-C	1-D	2-A	2-B	2-C	2-D	3-A	3-B	3-C	3-D
Flow Rate (ml/min)												
Temperature (°C)												

APPENDIX A-1 (Cont.)

RECORD OF ADULT FISH SURVIVAL AND WATER QUALITY CHARACTERISTICS

STUDY: _____

Date: _____ **Initials:** _____

Water Quality Characteristics (cont.)

	1-A	1-B	1-C	1-D	2-A	2-B	2-C	2-D	3-A	3-B	3-C	3-D
Dissolved Oxygen (mg/L)												
% Sat												
Alkalinity (mg/L $CaCO_3$)												
Hardness (mg/L $CaCO_3$)												
pH												
Un-ionized Ammonia (µg/L)												
Total Organic Carbon (mg/L)												

Comments:

APPENDIX A-2

RECORD OF SPAWNING ACTIVITY

STUDY: _____

Treatment _____ **Replicate** _____

Test Day	Init	Date	Total # Spawns	# Eggs/ Spawn	# Fertile/ Spawn	Develop. Stage	Incubation ID/ Comments
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							

APPENDIX A-3

FATHEAD MINNOW SAMPLE RECORD

STUDY:_____

Date:_____
Initials:_____

Treatment _____ Replicate _____	Tissue Sample/ID Number/Comments
Sex M / F Body Weight _____g Gonad Weight _____mg	Plasma ID _____ Gonad Histology ID _____ Aromatase ID _____ Receptor ID _____ Liver ID _____ Brain ID _____ Body ID _____
Secondary Sex Characteristics/Comments Ovipositor P / A _____ Tubercles P / A _____ Dorsal Pad P / A _____ Coloration P / A _____	
Other comments:	

Treatment _____ Replicate _____	Tissue Sample/ID Number/Comments
Sex M / F Body Weight _____g Gonad Weight _____mg	Plasma ID _____ Gonad Histology ID _____ Aromatase ID _____ Receptor ID _____ Liver ID _____ Brain ID _____ Body ID _____
Secondary Sex Characteristics/Comments Ovipositor P / A _____ Tubercles P / A _____ Dorsal Pad P / A _____ Coloration P / A _____	
Other comments:	

APPENDIX A-4

REPORT FORM FOR THE 21-D REPRODUCTION TEST

Chemical: _____ Test Start Date: _____
Investigator: _____ Test Termination Date: _____

TEST CHEMICAL

CAS Number: _____ MF: _____ MWT: _____
Chemical Source: _____ K_{ow}: _____ Vapor Pressure: _____
Lot Number: _____ Water Solubility: _____ Purity: _____

Acute Toxicity: _____ Reference: _____

Chronic Toxicity: _____ Reference: _____

Range-finding Test Information: _____

TEST ANIMALS

Source: _____ Age at Test Start: _____
Condition: _____ Pre-Exposure Reproduction: _____
Reference Toxicity Test EDC Performance Standard Test
Chemical: _____ Chemical: _____
Results: _____ Results: _____
Date: _____ Date: _____

TEST CONDITIONS

Dilution Water Source: _____ Dilution Water Treatment: _____

Water Temperature (°C)	Mean (± SD) _____	Minimum: _____	Maximum: _____
Dissolved Oxygen (mg/L)	Mean (± SD) _____	Minimum: _____	Maximum: _____
pH	Mean (± SD) _____	Minimum: _____	Maximum: _____
Hardness (mg/L $CaCO_3$)	Mean (± SD) _____	Minimum: _____	Maximum: _____
Alkalinity (mg/L$CaCO_3$)	Mean (± SD) _____	Minimum: _____	Maximum: _____
TOC (mg/L)	Mean (± SD) _____	Minimum: _____	Maximum: _____
Un-ionized Ammonia (µg/L)	Mean (± SD) _____	Minimum: _____	Maximum: _____

Illumination Source: _____ Photoperiod: _____
Light Intensity: _____
Exposure Chamber Dimensions: _____ Water Volume: _____
Flow Rate: _____
Number of Treatments: _____ Number of Replicates/Treatment: _____
Number of Females/Chamber: _____ Number of Males/Chamber: _____
Number of Substrates/Chamber: _____ Composition of Substrate: _____
Feeding Regime: _____

APPENDIX A-4 (Cont.)

Type of Chemical Administration: _____
(e.g. aqueous without solvent, aqueous with solvent, dietary, i.p. injection)

If aqueous administration was used, specify:
Method of Stock Generation: _____ Mean Concentration of Stock(s): _____

If solvent carrier was used, specify:
Solvent: _____ Maximum Solvent Concentration: _____

If dietary exposure was used, specify:
Food: _____ Method of Chemical Incorporation: _____
Source of Food: _____ Food Lot Number: _____

If i.p. exposure was used, specify:
Carrier Solvent: _____ Carrier Solvent Concentration: _____
Volume of Carrier Solvent Injected: _____

TEST CHEMICAL CONCENTRATIONS				
Date	Control	Control*	Treatment 1	Treatment 2
Nominal:				
Measured				
Date:				
Date:				
Date:				
Date:				
Date:				
Date:				
Mean (± SD)				
Range				

Mean (± SD) % Recovery of Spiked Samples: _____ (N = _____)
Mean (± SD) Repeatability of Duplicate Analysis: _____ (N = _____)
Analytical Methodology _____
Method Detection/Quantitation Limits _____

* Solvent control or sham-injected control.

APPENDIX A-4 (Cont.)

TEST RESULTS (Mean ± SD)[a]				
Primary Endpoints[b]	Control	Control[c]	Treatment 1	Treatment 2
Adult Survival (%)				
Reproductive Behavior (specify)				
Secondary Sex Characteristics (specify ovipositor, tubercles, dorsal pad, coloration)				
GSI (%)				
Gonadal Histopathology (specify)				
Plasma vitellogenin (mg/ml)				
Plasma sex steroids (specify β-estradiol, testosterone, 11-ketotestosterone) (ng/ml)				
Fecundity (specify total eggs, number of spawns/female, number of eggs/spawn)				
Fertility (%)				
Optional Endpoints				
Embryo Hatch (%)				
Larval Survival (%)				
Larval Morphology (specify)				

[a] Statistical significance, as well as the statistical tests used, level of significance, and sample size should be indicated.
[b] Adult survival, reproductive behavior, secondary sex characteristics, GSI, gonadal histopathology, plasma vitellogenin and sex steroids must be reported on a sex-specific basis.
[c] Solvent control or sham-injected control.

General Remarks

APPENDIX B

MEASUREMENT OF PLASMA VITELLOGENIN IN FATHEAD MINNOWS BY COMPETITIVE ELISA

OPERATING PROCEDURE

1. SCOPE AND APPLICATION

This procedure is used to at the MED laboratory to determine the concentration of vitellogenin (Vtg) in plasma of fathead minnows (*Pimephales promelas*). Other validated techniques can be used to measure Vtg; the following is included as one option. The level of Vtg present is indicative of the presence of estradiol, or other estrogen-like compounds, which can bind the estrogen receptor and induce the synthesis of Vtg. The ability to detect Vtg in male fish is particularly significant since Vtg is normally undetectable or present at extremely low levels compared to female fish. The Vtg is detected using antibodies in an ELISA (enzyme linked immunosorbent assay) procedure. The detection limit for Vtg is routinely 2 µg per ml of plasma when using a 1:300 dilution, which is typical for most males.

2. SUMMARY OF METHOD

The assay is performed in 96-well ELISA plates that have been coated with purified fathead minnow Vtg. After blocking the unbound sites with goat serum, the plate is washed to remove any unbound material. The samples being assayed, a range of standards prepared with purified fathead minnow Vtg, and the appropriate blanks and controls are incubated with an antibody specific to fathead minnow Vtg. After this incubation, the samples are added to duplicate wells of the plate. The plate is incubated again, during which there is a competition for the antibody between the Vtg bound on the surface of the wells and the Vtg present in the samples and standards. This is followed by another wash which removes all the primary antibody which is not bound to the Vtg on the plate. The primary antibody that is bound to the plate is detected by incubation with a secondary antibody which is conjugated to horseradish peroxidase. After washing the plate, the activity of the enzyme is measured, which is inversely proportional to how much Vtg was present in the sample.

2.1 Definitions

Vtg = Vitellogenin
ELISA = Enzyme-linked Immunosorbent Assay
FHM = Fathead Minnow
1° Ab = Primary Antibody
2° Ab = Secondary Antibody
NGS = Normal Goat Serum
BSA = Bovine Serum Albumin

IAP = Interassay Pool Sample
PBS = Phosphate Buffered Saline
PBST = Phosphate Buffered Saline with 0.05 % Tween-20 (same as wash buffer)
NSB = Nonspecific Binding
B_0 = Maximum Binding
B = Binding

2.2 Health and Safety Warnings

Follow routine laboratory safety precautions such as wearing lab coats and safety glasses. In addition, wear latex or nitrile gloves when working with phosphoric or hydrochloric acid solutions to prevent severe damage by contact with skin.

2.3 Cautions

Vtg is a very labile protein; keep on ice and store at -20°C when not in use. Store plasma samples and 1° Ab at -80°C. Store 2° Ab at -20°C. The 10X PBS buffer can be stored for several months at 4°C; discard if the solution becomes cloudy.

3. PERSONNEL QUALIFICATIONS

Those who perform this assay should be trained and adept at performing sample dilutions and working with microliter and multichannel pipets. Accurate pipeting is the most important skill required to obtain quality results.

4. MATERIALS AND METHODS

Equipment:

Centrifuge for hematocrit tubes (e.g., Adams MHCT II)
Centrifuge for microcentrifuge tubes (e.g., Eppendorf 5417R)
Speed-Vac for lyophilizing aprotinin (e.g., Savant SVC100)
Water bath at 37°C
Multi-tube vortexer (e.g., VWR)
Vortex mixer
96-well plate reader (e.g., Bio-Rad 3550)
Multichannel (8 or 12) pipet capable of delivering volumes in the 100 to 200 µl range
Adjustable pipets for 0.5 to 10 µl, 10 to 1000 µl, and 1 to 10 ml range
Repeater pipet for delivering 300 µl
Computer spreadsheet software (e.g., Lotus 123, version 9.5)
Computer graphing software (e.g., SlideWrite, version 5)

Materials:

Heparinized microhematocrit tubes (e.g., Oxford 8889-301209)
0.6 ml microcentrifuge tubes
Aprotinin (e.g., Sigma A-6279)
Pipet tips
Parafilm
Air-tight container (e.g. Tupperware)
15 and 50 ml disposable centrifuge tubes
12 x 75 mm disposable glass test tubes
96 well EIA plates (e.g., ICN 76-381-04)
Purified FHM Vtg (Obtained from Dr. Nancy Denslow, University of Florida)
1° Ab for FHM Vtg (Obtained from Dr. Louise Parks, EPA)
Peroxidase-conjugated anti-rabbit 2° Ab (e.g., Bio-Rad 172-1019)
Normal goat serum (NGS) (e.g., Chemicon S26-100ml)
TMB peroxidase substrates (e.g., KPL 50-76-00)
1 *M* H_3PO_4
1 *M* HCl
Coating Buffer 50 *mM* carbonate buffer, pH 9.6:
- 1.26g $NaHCO_3$
- 0.68g Na_2CO_3
- 428 ml of deionized water

10X PBS (0.1 *M* phosphate, 1.5 *M* NaCl):
- 0.83g monobasic sodium phosphate (monohydrate)
- 20.1g dibasic sodium phosphate (heptahydrate)
- 71g NaCl
- 810 ml of deionized water

Wash buffer (PBST):
- 100 ml of 10X PBS
- 900 ml of deionized water
- 0.5 ml Tween-20 (e.g., Sigma P-7949)
- Adjust pH to 7.3 with 1 *M* HCl

Blocking solution (5% NGS in Coating Buffer)
Assay buffer (PBST + 2.5% NGS)

4.1 Sample Collection

Collect blood from severed caudal artery/vein with a heparinized microhematocrit tube and place on ice or other cooling device. It is important to keep the blood cold to minimize Vtg degradation. Centrifuge for 3 min, score tube and expel plasma into 0.6 ml microcentrifuge tubes containing 0.13 units of lyophilized aprotinin. (Prepare these tubes in advance by adding the appropriate amount of aprotinin solution, freezing, and

lyophilizing in a speed-vac at low heat for approximately 20 min, or until no liquid remains). Mix samples gently and centrifuge briefly to collect the contents at the bottom of the tube. Store at -80°C until analysis.

4.2 Analysis Procedure

The instructions describe the method for processing one plate at a time. Two or three plates may be processed simultaneously, but standards and appropriate blanks/controls must be included in each plate.

4.2.1 Coat Plate

Dilute purified FHM Vtg to 0.56 µg/ml in coating buffer. (The actual concentration is determined empirically and may change with different batches of purified VTG). Add 200 µl to each well of the plate. Cover the plate with parafilm and place in a sealed container (e.g., Tupperware) along with some moistened paper towels. Incubate the container overnight at 4°C (or for 2 h at 37°C).

4.2.2 Block Plate

Shake out the coating solution and pat the plate dry on absorbent paper. Add 350 µl of blocking solution to each well, cover with parafilm, place in a container like above, and incubate for 2 h at 37°C (or overnight at 4°C if the plate will not be used that day; the plate should be used within a few days).

4.2.3 Prepare Standards

Dilute purified Vtg standard (concentration determined by supplier using the common Bradford Method and comparison to BSA) to 2000 ng/ml and 750 ng/ml in 2 ml of assay buffer. Mix one part (0.5 ml) of each of these with three parts (1.5 ml) of assay buffer to yield 500 and 188 ng/ml concentrations, respectively. Prepare additional dilutions in a similar manner until the entire range of desired standards is obtained (2000, 750, 500, 188, 125, 47, 31, 12, 8, and 3 ng/ml).

4.2.4 Dilute Samples

Because of normal variation in fathead minnow Vtg concentrations, the following should be considered a general guideline. For a typical male with little or no Vtg, combine 1 µl of plasma and 300 µl of assay buffer directly in the test tube that will be used for incubation with the 1° Ab. For females or males that may have high Vtg levels, additional dilution is usually required. Dilute an initial 300-fold dilution further by combining 5 µl with 2 ml of assay buffer giving a total dilution of 120,000-fold. One or two additional 1:2 dilutions (1 part diluted sample to 1 part assay buffer)

of the 120K dilution are usually necessary to achieve a Vtg sample concentration within the reliable range of the standards (approximately 10 to 200 ng/ml). If there are enough sample wells available, it is advantageous to analyze samples of female plasma (or male plasma if Vtg is highly elevated) at the 120K, 240K, and 480K dilution, but just one of the dilutions is generally sufficient if the result is within the reliable range of the standards. Put 300 µl of each dilution in a test tube for incubation with the 1° Ab.

4.2.5 Prepare Interassay Pool (IAP) Sample

An interassay pool sample should be analyzed as a quality control procedure each time an assay is performed. Sufficient plasma for this sample can be obtained by combining unused plasma remaining from previously analyzed samples into a large pool, separating into aliquots of a few µl, and storing at -80°C. Each time an assay is performed, remove one of these aliquots and dilute as necessary to get a result in the reliable range of the standards.

4.2.6 Prepare Maximum (B_0) and Non-specific Binding (NSB) Samples

Prepare B_0 by putting 300 µl of assay buffer in a test tube. Prepare NSB by putting 600 µl of assay buffer in a test tube.

4.2.7 Incubate with Primary Antibody

The 1° Ab must be diluted to give the desired response. The antibody has previously been diluted 1:2 with an equal volume of normal male FHM plasma. On the day of use, dilute the antibody 11K-fold (1.8 µl to 20 ml) in assay buffer. This dilution is determined empirically and will likely change from batch to batch. After diluting the 1° Ab, add 300 µl to each sample/standard tube (except the NSB sample) using the repeater pipet. Mix the tubes briefly and incubate for 1 h at 37°C.

4.2.8 Wash Plate

Approximately 5 min before the end of the 1° Ab incubation, shake out the blocking solution and pat the plate dry on absorbent paper. Then, using a multichannel pipet, fill the wells with 400 µl of wash buffer, remove by shaking, and pat dry. Repeat this procedure two times.

4.2.9 Load Samples on Plate

At the end of the 1° Ab incubation, remove the tubes from the water bath and slightly vortex. Remove two-200 µl aliquots and add to each of the duplicate wells of the 96-well plate. A template identifying sample location is useful. After the samples are

loaded, cover the plate with parafilm and incubate the container for 1 h at 37°C as before.

4.2.10 Incubate with Secondary Antibody

Before the first use, divide the 2° Ab into aliquots of 5-10 µl and store at -20°C. Fifteen min before the end of the sample incubation, remove an aliquot of the 2° Ab from the freezer and dilute 40K (1 µl to 40 ml) with assay buffer. At the end of the sample incubation, wash the plate as described above (4.2.8). Then add 200 µl of the 2° Ab solution to each well. Cover the plate with parafilm and incubate the container for 1 h at 37°C as before.

4.2.11 Enzymatic Conversion of Substrate to Colored Compound

Approximately 15 min before the end of the 2° Ab incubation, turn on the plate reader. Measure 7 ml each of the two peroxidase substrate solutions from the TMB kit (this is sufficient for one 96-well plate) and allow solutions to come to room temperature by placing them in a water bath. At the end of the 2° Ab incubation, wash the plate as described above (4.2.8). Mix the two substrate solutions and add 100 µl of the mixture to each well of the plate using the multichannel pipet. Do this in a timed manner. After 5 to 10 min, the color in the wells will change to blue. Stop the reaction by adding 100 µl of 1 *M* phosphoric acid, which changes the color in the wells to intense yellow. The exact time of reaction is not critical, but each well should react for the same amount of time. The goal is to achieve a net absorbance of 1.5 to 2.0 in the wells with the most color.

4.2.12 Read Absorbance with Plate Reader

Set the plate reader to measure at 450 nm with the lower limit at 0.000 and the upper limit at 3.000.

4.5 Troubleshooting

Normal results will produce a sigmoidal curve (see Section 5) with the steepest part between 10 and 200 ng/ml. When changing to a new source of purified Vtg or 1° Ab, the dilutions may have to be adjusted to obtain the proper curve. Lowering the concentration of the coating Vtg will shift the curve to the left, or a more sensitive area. Likewise, a more dilute 1° Ab will also make the assay more sensitive. While a more sensitive assay is generally good, performing the assay in this range will narrow the range of concentrations that can be detected. Due to the normal variation between fish, having a narrow effective range of the standards would not necessarily be advantageous. Also, making the assay more sensitive decreases the net absorbance response, which will adversely affect detection.

5. DATA ACQUISITION, CALCULATIONS, AND DATA REDUCTION

The absorbance values from the plate reader are obtained on a printout and entered into a spreadsheet (e.g., Lotus 123 version 9.5) for performing the necessary calculations.

5.1 Calculate B/B_0

Calculate the B/B_0 for each sample and standard. Divide the absorbance value by the absorbance of the B_0 sample, after the NSB value has been subtracted. Determine the B/B_0 for each duplicate well of the standards individually; for the samples, determine the mean B/B_0 of the two duplicate wells.

5.2 Obtain Standard Curve

Generate the sigmoidal standard curve using graphing software (e.g. SlideWrite 5). Plot the B/B_0 of the standards against the amount of standard present on a log scale.

5.3 Determine Corresponding Amount of Vtg in Samples from B/B_0 Values

Use the sample B/B_0 values from the spreadsheet to generate extrapolated values of Vtg present in the sample using the graphing software.

5.4 Calculate Amount of Vtg per ml of Plasma

Transfer the extrapolated sample Vtg values to the spreadsheet. Multiply these values by the appropriate dilution factor and convert to mg/ml by dividing by 1,000,000, giving mg Vtg per ml of plasma.

5.5 Determine Minimum Detection Limits

Often, particularly in normal males, it may be difficult to determine whether the result is truly zero or slightly greater than zero. If this is the case, use the 95% confidence limit (calculated with spreadsheet software) to determine if the result should be reported as zero or as greater than zero. If the sample result is within the confidence interval of the zero standard (NSB), report the result as zero. The minimum detection level is the lowest standard which is consistently different than the zero standard; that is, the two confidence limits do not overlap. For any sample result which is within the confidence limit of the minimum detection level or above, report the calculated value. If the sample result falls between the zero standard and the minimum detection level, report one-half of the minimum detection level.

5.6 Data Management and Records Management

Attach printed versions of any computer generated data in a bound record book along with any other necessary descriptions of how the data were generated.

6. QUALITY CONTROL AND QUALITY ASSURANCE SECTION

6.1 Sigmoidal Shaped Standard Curve

The standard curve should have a typical sigmoidal appearance, flattening out at both the low and high end of the standard range. The coefficient of determination for the curve (determined from graphing software) should be around 0.99. If it is sufficiently different, the plot should be examined to determine if there are potential outliers. Also, the sensitivity should not differ from previous analyses. This would be indicated by a change in the steeper portion of the curve relative to the standards.

6.2 Low NSB Value

The NSB should be approximately 0.06, or less.

6.3 Consistent IAP Results.

The IAP result should not differ from previous analyses, or drift slowly over several experiments.

6.4 Dependable Pipets

Check the operation of the pipets periodically for accuracy and precision. Fix or replace any pipets not operating correctly.

6.5 Duplicate Agreement

Duplicate agreement should be within 10%.

6.6 Agreement with Different Dilutions

When multiple dilutions of the same sample are assayed, the final results should be reasonably similar. However, the most accurate results are obtained when the B/B_0 values are approximately 50%.

7. REFERENCES

Parks LG, Cheek AO, Denslow ND, Heppell SA, McLachlan JA, LeBlanc GA, Sullivan CV. 1999. Fathead minnow (*Pimephales promelas*) vitellogenin: purification, characterization, and quantitative immunoassay for the detection of estrogenic compounds. *Comp Biochem Physiol* C 123:113-125.

Jensen KM, Korte JJ, Kahl MD, Pasha MS, Ankley GT. 2001. Aspects of basic reproductive biology and endocrinology in the fathead minnow (*Pimephales promelas*). *Comp Biochem Physiol* C 128: 127-141.

Korte JJ, Kahl MD, Jensen KM, Pasha MS, Parks LG, LeBlanc GA, Ankley GT. 2000. Fathead minnow vitellogenin: complementary DNA sequence and messenger RNA and protein expression after 17β-estradiol treatment. *Environ Toxicol Chem* 19:972-981.

APPENDIX C

DETERMINATION OF PLASMA STEROID CONCENTRATIONS IN FATHEAD MINNOWS BY RIA

OPERATING PROCEDURE

1. SCOPE AND APPLICATION

This procedure is used at the MED laboratory to measure the concentration of sex steroids in plasma of fathead minnows (*Pimephales promelas*) by radioimmunoassay (RIA). With relatively minor changes in reagents, it is used to measure β-estradiol (E2), testosterone (T), and 11-ketotestosterone (11-KT). Other validated techniques can be used to measure these steroids; the following is included as one option. The levels of steroids determined from these assays are used to assess the endocrine status of fish used in research on endocrine disrupting chemicals (EDCs). Typically, detection limits are about 0.4 ng/ml of plasma when performing the assay as described herein.

2. SUMMARY OF METHOD

This assay is performed on ether-extracts of plasma to minimize complications from interfering substances. The extract is incubated with a small amount of radioactive steroid and an antibody which recognizes the steroid of interest. The unlabeled steroid present in the sample and the radioactive steroid compete to bind to the antibody. The more steroid present in the sample, the less radioactive steroid the antibody can bind. The addition of a charcoal solution binds all of the unbound steroid. After centrifugation to remove the charcoal, the radioactivity is measured in an aliquot of the supernatant fluid.

2.1 Definitions

RIA = Radioimmunoassay
E2 = β-Estradiol
T = Testosterone
11-KT = 11-Ketotestosterone
BSA = Bovine Serum Albumin
PBS = Phosphate Buffered Saline
NSB = Nonspecific Binding
B_0 = Maximum Binding
B = Binding
^3H = Tritium or Tritiated
IAP = Interassay Pool

2.2 Health and Safety Warnings

2.2.1 General

Follow routine laboratory safety precautions such as wearing lab coats and safety glasses. In addition, wear latex or nitrile gloves when working with sodium hydroxide, hydrochloric acid, ethanol, and scintillation cocktail to prevent severe damage by contact with skin.

2.2.2 Working with Radioactivity

Small amounts of ^{3}H are used in these assays. Wear gloves when handling ^{3}H. Training in safe handling of radioactive materials is required by the NRC (Nuclear Regulatory Commission).

2.2.3 Ether

Ether is a very volatile, extremely flammable, toxic compound and should be used in a fume hood as much as possible. Cover tubes with parafilm when removing them from the hood. Wear gloves to prevent contact with skin.

2.2.4 Steroids

These steroids are hormones and thus have activity in humans. They also may possess carcinogenic and mutagenic activities. It is important not to come in contact, ingest, or breathe these compounds. Wear gloves to prevent contact with skin.

2.3 Cautions

Care should be taken not to spill radioactive solutions, especially when working with the stock vials. Work with absorbent toweling, preferably with plastic backing, under all work areas. If a spill should occur, it must be cleaned up and the spill area decontaminated in accordance with NRC requirements.

3. PERSONNEL QUALIFICATIONS

Those who perform this assay should have completed a radiation safety course. Personnel should also be trained and accomplished at making sample dilutions and working with microliter volumes.

4. MATERIALS AND METHODS

Equipment:

Scintillation counter (e.g., Packard 2500 TR)
Centrifuge (e.g., Jouan CR412)
Vortex mixer
Multi-tube vortexer (e.g., VWR)
Water bath
Pipets for 0.5 to 10 µl, 10 to 1000 µl, and 1 to 10 ml range
Repeater pipet for delivering 100 µl
Repeater diluter for removing aliquot/adding scintillation fluid (e.g., Brinkman)
Bottle-top dispenser for dispensing ether (e.g., Brinkman)
Computer spreadsheet software (e.g., Lotus 123, version 9.5)
Computer graphing software (e.g., SlideWrite, version 5)

Materials:

Heparinized microhematocrit tubes (e.g., Oxford 8889-301209)
0.6 ml microcentrifuge tubes
Parafilm
Disposable 12 X 75 mm borosilicate tubes
7 ml glass scintillation vials
Pipet tips
β-Estradiol (e.g., Sigma E-2758)
Testosterone (e.g., Sigma T-1500)
11-ketotestosterone (e.g., Sigma K-8250)
Tritiated E2 (e.g., Amersham TRK 587)
Tritiated T (e.g., Amersham TRK 921)
Tritiated 11-KT (e.g., Amersham TRQ 7903, Custom Synthesis)
Antibody (antisera) for E2 (e.g., Endocrine Sciences E26-47)
Antibody (antisera) for T (e.g., Endocrine Sciences T3-125)
Antibody (antisera) 11-KT (e.g., Biosense Laboratories)
Scintillation cocktail (e.g., Fisher Scintiverse SX18-4)
Ethyl ether
Parafilm
Ethanol
0.1 *M* Phosphate Buffered Saline (PBS), pH 7.6
  11.36 g Na_2HPO_4
  3.12 g $NaH_2PO4{\cdot}2H_2O$
  8.76 g NaCl
  Deionized water to ~950 ml,
  Adjust pH to 7.6 with solid NaOH and 1 *M* NaOH
  Deionized water to1000 ml

Assay buffer, or 0.01 *M* PBS, pH 7.4, with 1% BSA
100 ml of 0.1 *M* PBS, pH 7.6
900 ml of deionized water
7.9 g NaCl
Adjust pH to 7.4 with 1 *N* HCl
Add 10 g BSA (e.g., Sigma A-7888)
Dextran-coated charcoal solution
1.5 g Activated charcoal (e.g., Sigma C-5260)
0.15 g Dextran (e.g., Sigma D-4751)
300 ml of 0.1 M PBS, pH 7.6
Mix at 4°C overnight before use. Keep at 4°C until just before using.
Steroid Stock Solutions
A) Dissolve 10.0 mg steroid in 10.0 ml of ethanol
B) Mix 100 µl of (A) with 900 µl of ethanol
C) Mix 10 µl of (B) with 990 µl of ethanol, aliquot at ~ 20 µl/vial and store at -80°C

4.1 Sample Collection

Collect blood from severed caudal artery/vein with a heparinized microhematocrit tube and place on ice or other cooling device. It is important to keep the blood cold since these samples will often be used for vitellogenin (Vtg) analyses also. Centrifuge for 3 min, then score tube and expel plasma into 0.6 ml microcentrifuge tubes containing 0.13 units of lyophilized aprotinin. (Prepare these tubes in advance by adding the appropriate amount of aprotinin solution, freezing, and lyophilizing in a speed-vac at low heat for approximately 20 min, or until no liquid remains. If Vtg will not be analyzed, tubes containing aprotinin are not necessary). Mix samples gently and centrifuge briefly to collect the contents at the bottom of the tube. Store at -80°C until analysis.

4.2 Ether Extraction and Reconstitution of Plasma

A convenient number of samples to extract is approximately 24. In addition, an interassay pool sample should be analyzed as a quality control procedure each time an assay is performed. (Sufficient plasma for this sample can be obtained by combining unused plasma remaining from previously analyzed samples into a large pool, separating into aliquots, and storing at -80°C). Do not prepare more RIA tubes for any one steroid analysis than the centrifuge can accommodate. It is important that all of the tubes for a particular steroid analysis be centrifuged at the same time.

4.2.1 Measure Volume of Plasma

Thaw plasma and, if possible, remove 6 µl for every determination that is to be made. For example, in males it is often desirable to measure E2, T, and 11-KT, and since it is desirable to perform duplicate analyses, it would be optimal to have 36 µl of

plasma. Often it is not possible to collect 36 µl of plasma; if this is the case, priority must be assigned to individual steroids for analyses. Record the volume of plasma removed from each sample and placed in the test tube for extraction so that it can be properly reconstituted.

4.2.2 Add PBS

Add 150 µl of 0.1 M PBS to each tube. This is essentially to give some volume to the aqueous phase, although pH and ionic strength may also be important.

4.2.3 Add Ether

Use the bottle-top dispenser, if available, to deliver 1.5 ml of ether to each tube. Lay a sheet of parafilm over all the tubes and vortex for 1 min using the multi-tube vortexer. Adjust speed to highest possible without the ether contacting the parafilm. If ether contacts parafilm, some waxes may be extracted which will interfere with redissolving the steroids.

4.2.4 Place Samples in Freezer

Wait 1 min after vortexing to allow the phases to completely separate. Place the rack in a -80°C freezer for 10 min to freeze the lower aqueous phase.

4.2.5 Decant Ether

Pour the upper ether layer into another clean, labeled test tube. Try to get as much of the ether as possible, but work quickly enough that all the aqueous phases remain frozen.

4.2.6 Repeat Extraction

Thaw the aqueous phase and repeat steps 4.2.3 to 4.2.5. Decant the ether from the second extraction into the same tube as the first extraction.

4.2.7 Evaporate Ether

Leave the tubes in the fume hood overnight to evaporate all the ether.

4.2.8 Reconstitute Samples

Add 120 µl of assay buffer to each tube for each 6 µl-aliquot removed in 4.2.1. Lay a sheet of parafilm over all the tubes and vortex for 1 min using the multi-tube vortexer. (It may be necessary to manually reconstitute the samples by rinsing the

walls of the extraction tube with assay buffer when the volume is too low for the vortexer to wet the sides of the tube.) If there is solution clinging to the sides of the tube, it may be necessary to centrifuge briefly to collect the contents at the bottom.

4.3 Incubate Samples with Antibody and Radioactive-tracer

4.3.1 Prepare Standards

Thaw an aliquot of Standard Solution C for each steroid that will be measured. Prepare a 10 ng/ml standard; add 10 µl of Standard Solution C to 990 µl of assay buffer in a test tube and vortex to mix thoroughly. Complete a series of 1:2 (1 part to 1 part) dilutions with assay buffer (e.g. 0.5 ml) to obtain standards of 5, 2.5, 1.25, 0.625, 0.312, 0.156, 0.078, 0.039, 0.020, and 0.010 ng/ml.

4.3.2 Prepare Antibody Solution

The following should be considered a general guideline; specific concentrations may change with different batches of antibodies. Initial experimentation with new batches should attempt to determine the appropriate dilution resulting good sensitivity between the 0.1 to 1.0 ng/ml standard range.

4.3.2.1 Estradiol

Resuspend as directed by the supplier, aliquot in approximately 100 µl portions, and store at -80°C. On the day of use, thaw an aliquot and dilute 70 µl with 10 ml of assay buffer. Mix gently, but thoroughly.

4.3.2.2 Testosterone

Resuspend as directed by the supplier, aliquot in approximately 100 µl portions, and store at -80°C. On the day of use, thaw an aliquot and dilute 110 µl with 8 ml of assay buffer. Mix gently, but thoroughly.

4.3.2.3 11-Ketotestosterone

This antibody was previously aliquoted and stored at -80°C. On the day of use, thaw an aliquot and dilute 1 µl with 12 ml assay buffer. (The aliquot of antibody can be frozen again and reused several times before starting a new aliquot). Mix gently, but thoroughly.

4.3.3 Prepare Tritiated Tracer

For each steroid being measured, dilute 1 µl of the stock to10 ml with assay buffer. (The actual concentration is determined empirically and may change with different batches of isotope). There should be 5000 - 6000 cpm in the 0.5 ml aliquot that is counted from the Total tube.

4.3.4 Incubate Sample/Standards with Antibodies and Tracers

Analyze standards, controls, and if possible, the samples in duplicate. Add 100 µl of reconstituted sample extract and standards to labeled test tubes. Add 100 µl of assay buffer to the B_0 tube and 200 µl assay buffer to Total and NSB tubes. Using the repeater pipet, add 100 µl of diluted antibody to all tubes except the Total and NSB tubes. Using the repeater pipet, add 100 µl of tracer solution to all tubes. Lay a sheet of parafilm over the tubes and mix gently using the multi-tube vortexer. Incubate at 25°C for 1.5 to 2 h. If only two steroids are being measured, they can usually incubate simultaneously, but if all three steroids are being measured, stagger the start time of the incubation by 45 min to 1 h for one of the steroids so that conflicts due to limited centrifuge space are avoided.

4.4 Add Charcoal Solution

At the end of the incubation, place the tubes in an ice-water bath for 15 min. Remove the charcoal solution from the refrigerator a few minutes before needed and stir vigorously to assure that a uniform suspension is obtained. While the charcoal solution is still being stirred, add 400 µl to all tubes except the Total tubes using the repeater pipet. Add 400 µl of 0.1 *M* PBS to Total tubes. Vortex gently and incubate again in an ice-water bath for 15 min. Centrifuge the tubes at 3000 rpm for 30 min at 4°C.

4.5 Load Scintillation Vials

Adjust the repeater-diluter to remove 0.5 ml of the supernatant and deliver 5 ml of scintillation cocktail. Carefully place the inlet/outlet tube below the liquid surface in the tube and slowly remove 0.5 ml with the uptake mechanism, being careful not to disturb the charcoal at the bottom of the tube. Then, while holding the mechanism so that it does not release, put a scintillation vial under the inlet/outlet tube and expel the contents that were removed from the tube. Then initiate the delivery of the scintillation fluid. Place a cap on the vial immediately. When a whole rack of tubes has been processed, mix by shaking.

4.7 Determine Radioactivity

Count vials in a scintillation counter for 1 min using a program appropriate for tritium.

4.5 Troubleshooting

Normal results will produce a sigmoidal curve (see Section 5) with the steepest part between approximately 0.1 and 1 ng/ml. When changing to a new source of antibody, the dilutions may have to be adjusted to obtain the proper curve. Although the amount of steroid in the tracer solutions can also affect the curve, a more dilute Ab will make the assay more sensitive, shifting the curve to the left. While a more sensitive assay is generally good, performing the assay in this range will narrow the range of concentrations that can be detected. Due to the normal variation between fish, having a narrow effective range of the standards would not necessarily be advantageous. Also making the assay more sensitive decreases the net cpm response, which will adversely affect detection.

5. DATA ACQUISITION, CALCULATIONS, AND DATA REDUCTION

The cpm values from the counter are obtained on a printout and entered into a spreadsheet (e.g., Lotus 123) for performing the necessary calculations.

5.1 Calculate B/B_0

Calculate the B/B_0 for each sample and standard. Divide the cpm value by the cpm of the B_0 sample, after the NSB value has been subtracted. Determine the B/B_0 for each duplicate well of the standards; for the samples, determine the mean B/B_0 of the two duplicates.

5.2 Obtain Standard Curve

Generate the sigmoidal standard curve using graphing software (e.g., SlideWrite 5). Plot the B/B_0 of the standards against the amount of standard present on a log scale.

5.3 Determine Corresponding Amount of Steroid in Samples from B/B_0 Values

Use the sample B/B_0 values from the spreadsheet to determine the concentration of steroid present in the sample.

5.4 Calculate Amount of Steroid per ml of Plasma

Transfer the extrapolated steroid values to the spreadsheet. Multiply these values by 20 (dilution factor from 5 µl of plasma present in the 100 µl of reconstituted extracted) to give ng/ml plasma. Occasionally, an additional factor may be needed; for example, if steroid concentrations are expected to be very high, it may be sufficient to use only 25 or 50 µl of the extracted sample rather than 100 µl.

5.5 Determine Minimum Detection Limits

Often, particularly when measuring E2 in normal males, it may be difficult to determine whether the result is truly zero or slightly greater than zero. If this is the case, use the 95% confidence limit (calculated with spreadsheet software) to determine if the result should be reported as zero or as greater than zero. If the sample result is within the confidence interval of the zero standard (NSB), report the result as zero. The minimum detection level is the lowest standard which is consistently different than the zero standard; that is, the two confidence intervals do not overlap. For any sample result which is within the confidence limit of the minimum detection level or above, report the calculated value. If the sample result falls between the zero standard and the minimum detection level, report one-half of the minimum detection level.

5.6 Data Management and Records Management

Attach printed versions of any computer generated data in a bound record book along with any other necessary descriptions of how the data were generated.

6. QUALITY CONTROL AND QUALITY ASSURANCE SECTION

6.1 Sigmoidal Shaped Standard Curve

The standard curve should have the typical sigmoidal appearance, flattening out at both the low and high end of the standard range. The coefficient of determination for the curve (determined from graphing software) should be around 0.99. If it is sufficiently different, the plot should be examined to determine if there are potential outliers. Also, the sensitivity should not differ from previous analyses. This would be indicated by a change in the steeper portion of the curve relative to the standards.

6.2 Low NSB Value

The NSB value should generally be less than 400 to 600 cpm. The NSB values from a new radioactive stock solution usually start around 100 to 200 cpm, but can increase to 600 or 800 cpm. Replace the radioactive stock solution when the value becomes too high.

6.3 Consistent IAP Results.

The IAP result should not differ from previous analyses, or drift slowly over several experiments.

6.4 Dependable Pipets

Check the operation of the pipets periodically for accuracy and precision. Fix or replace any pipets not operating correctly.

6.5 Duplicate Agreement

Duplicates agreement should be within 10%.

6.6 Ratio of B_0 to the Total Counts

The maximum binding should be around 30 to 50% of the total counts in the system. Adding less antibody will lower this value and make the assay more sensitive. Likewise, increasing the antibody concentration will increase the B_0 and decrease sensitivity.

6.7 Check of Extraction Efficiency

Check the extraction efficiency periodically by spiking a sample with a known amount of steroid and measuring for recovery (especially if a drift in the IAP is observed). Extraction efficiency should be at least 85 to 90%.

7. REFERENCES

Kagawa H, Takano K, and Nagahama Y. 1981. Correlation of plasma estradiol-17 β and progesterone levels with ultrastructure and histochemistry of ovarian follicles in the white spotted char, *Salvelinus leucomaeni*. *Cell Tiss Res* 218: 315-329.

Jensen KM, Korte JJ, Kahl MD, Pasha MS, Ankley GT. 2001. Aspects of basic reproductive biology and endocrinology in the fathead minnow (*Pimephales promelas*). *Comp Biochem Physiol* C 128: 127-141.

APPENDIX D

HISTOLOGICAL TECHNIQUES FOR FATHEAD MINNOW GONADS

This appendix discusses two approaches to fixation and embedding appropriate for fathead minnow gonadal histology: a traditional paraffin-based approach and a more modern methacrylate-based procedure. Either technique is acceptable in the context of the test described in this document.

1. PARAFFIN-BASED HISTOLOGICAL PROCEDURE

1.1 FIXATIVE

1.1.1 Two fixatives have generally been used for paraffin-based studies of gonadal histology and histopathology in fathead minnows: (1) 10% neutral buffered formalin (McCormick et al. 1989; formula example: Roberts 1978), and (2) Bouin's fluid (Roberts 1978). Presumably Bouin's is used because it provides for rapid fixation (4-6 h) and strong subsequent tissue staining. However, intracellular substances, such as granules and inclusions, are often poorly preserved with this fixative (Kiernan 1990). Neutral buffered formalin gives better tissue preservation, but the gonads should ideally be fixed for 24 h or more.

1.2 EMBEDDING

1.2.1 Paraffin embedding of fathead minnow gonads can be accomplished by standard methods. An example of this procedure, using a tissue processor, is outlined below:

70% ethanol	1h
70% ethanol	1h
80% ethanol	1h
80% ethanol	1h
95% ethanol	1h
95% ethanol	1h
100% ethanol	1h
100% ethanol	1h
Xylene	1/2 h
Xylene	1/2 h
Paraffin	1h
Paraffin	1h
Paraffin	1h

1.3 SECTIONING AND STAINING

1.3.1 Standard sectioning and staining methods can be used with fathead minnow gonads. For example, sections may be cut at 5 µm and stained with hematoxylin and eosin (e.g., Kiernan

1990). Typically, the gonads are embedded longitudinally and sectioned in a step-wise fashion. For ovaries, a number of slides are made with one or two sections from 500 µm deep into the organ and one or two sections from 1000 µm deep. Testes are embedded and sectioned in a similar manner except that the sections are taken at 250 and 500 µm depths.

1.4 ADVANTAGES AND DISADVANTAGES

1.4.1 Advantages of utilizing paraffin techniques include: (1) the typical histology laboratory is set up to routinely process large numbers of samples, (2) more personnel are familiar with paraffin sectioning and staining procedures, and (3) paraffin sections have been used in several recent EDC studies. Disadvantages include: (1) thicker sections allow less detail to be resolved, (2) numerous artifacts (e.g., due to tissue shrinkage during processing), are present, and (3) tissue samples have to be archived in other fixatives for subsequent high-resolution electron microscopy.

2. JB-4 METHACRYLATE-BASED PROCEDURE

2.1 FIXATIVE

2.1.1 While the same fixatives (Bouin's, 10% neutral buffered formalin) used in paraffin procedures may be employed, the better tissue preservation afforded by formaldehyde-glutaraldehyde fixatives are preferable when embedding in methacrylate. Both traditional electron microscopic fixatives, e.g., 2.5% glutaraldehyde-2% formaldehyde in 0.1M phosphate buffer, or other fixatives, e.g., 1% glutaraldehyde-4% formaldehyde in 0.1M phosphate buffer (Jensen *et al.* 2001) work well with fathead minnow gonads. As with neutral buffered formalin, gonads should be fixed for at least 24 h prior to embedding.

2.2 EMBEDDING

2.2.1 Embedding of the relatively small fathead minnow gonads can be accomplished rapidly in JB-4 methacrylate compared to paraffin. A typical manual schedule for gonads (with tissue in vials on a rotator) is as follows:

25% ethanol	30 min
50% ethanol	30 min
75% ethanol	30 min
95% ethanol	30 min
100% ethanol	30 min

JB-4 solution A (catalyzed with 0.9 g catalyst/100 ml) 2 h
Embed (40 parts catalyzed solution A: 1 part solution B; prevent contact with air during polymerization)

2.3 SECTIONING AND STAINING

2.3.1 Gonads are embedded in the same longitudinal orientation as with paraffin blocks. They are also sectioned in the same step-wise manner as paraffin blocks, but at a thickness of 2 to 3 µm. Staining of methacrylate sections can be accomplished with most stains used in paraffin procedures, but with modifications. Examples of two modified procedures are given below.

Hematoxylin and eosin (with phloxin)

Stain sections for 30 to 45 min with filtered Harris hematoxylin
Rinse with distilled water
Dry on a hot plate
Stain cooled slides for 1 to 2 min in saturated aqueous eosin containing 0.25% phloxin
Rinse in distilled water, dry on a hot plate, and coverslip

Basic fuschin and methylene blue-azure A

Stock basic fuschin:
1 % basic fuschin in 50 % ethanol
Stock methylene blue-azure A in distilled water:
1 % azure A
1 % methylene blue
1 % borax
Staining procedure:
Dilute basic fuschin 1:4 to 1:12 or more with distilled water
Stain 10 to 20 sec and rinse with distilled water
Dilute methylene blue-azure A 1:2 to 1:4 or more with distilled water
Stain 10 to 20 sec, rinse with distilled water, dry and coverslip

2.4 ADVANTAGES AND DISADVANTAGES

2.4.1 Advantages of JB-4 methacrylate technique include: (1) better fixation, less solvent-related extraction, and thinner sections allow for greater resolution of tissue and cellular details compared to paraffin sections (the superior resolution afforded by methacrylate sections may obviate the need for electron microscopy in many cases), (2) tissues embedded in methacrylate experience little shrinkage while those embedded in paraffin shrink 20% or more, (3) cutting distortion, an artifact related to tissue compression during sectioning, is greatly reduced in thin methacrylate sections, (4) staining of methacrylate sections, for example, with hematoxylin and eosin, is far simpler than with paraffin sections, requiring no embedment removal or sequential hydration and dehydration steps, (5) when electron microscopy is required, the archived tissue is already in a suitable (recommended) fixative, and (6) methacrylate sections have been used in several recent EDC studies. Disadvantages include: (1) some laboratories lack the facilities for routine methacrylate procedures, and (2) staining procedures must be modified for methacrylate sections.

3. RECOMMENDATIONS

3.1 Where logistically possible, methacrylate embedding of tissues fixed for at least 24 h in formaldehyde-glutaraldehyde is recommended. The greater resolution and fewer artifacts seen in methacrylate sections make it worthwhile to establish a methacrylate facility in histopathology laboratories where EDC studies are performed. While most stains can be used with methacrylate, the staining procedure and stain ingredients may need to be modified. For example, when using thin 1-3 µm methacrylate sections, the standard tissue stain eosin may be combined with 0.1-0.5 % phloxin to improve color saturation.

4. REFERENCES

Jensen KM, Korte JJ, Kahl MD, Pasha MS, Ankley GT. 2001. Aspects of basic reproductive biology and endocrinology in the fathead minnow (*Pimephales promelas*). *Comp Biochem Physiol* C 128:127-141.

Kiernan JA. 1990. *Histological and Histochemical Methods: Theory and Practice*. 2nd ed. Pergamon Press, New York, NY, USA. 433 pp.

McCormick JM, Stokes GN, Hermanutz RO. 1989. Oocyte atresia and reproductive success in fathead minnows (*Pimephales promelas*) exposed to acidified hardwater environments. *Arch Environ Contam Toxicol* 18:207-214.

Roberts RJ. 1978. *Fish Pathology*. Bailliere Tindall, London, England. 318 pp.

APPENDIX E

HISTOLOGICAL EVALUATION OF FATHEAD MINNOW GONADS

1. NORMAL GONADAL HISTOLOGY IN REPRODUCTIVELY- MATURE FEMALES

1.1 GENERAL STRUCTURE

1.1.1 The ovaries are paired organs that, when mature, occupy much of the abdominal cavity ventral to the swim bladder (Grizzle 1979). They are suspended from the swim bladder by a mesentery, the mesovarium. The mesovarium is continuous with the peritoneum that covers the ovary proper. The peritoneum is composed of a mesothelium and an underlying layer of connective tissue, the tunica albuginea. In fathead minnows, rodlet cells are found among the squamous cells of the mesothelium, and eosinophilic granular cells, melanocytes, smooth muscle, blood vessels, and nerves are common within the connective tissue layer. The visceral border of the tunica albuginea is lined with an epithelium that is squamous in some regions and columnar in others. The columnar cells appear to be ciliated. Ovigerous lamellae extend from the tunica albuginea toward the center of the ovary, dividing it into lobules that contain the oogonia and developing oocytes (Grizzle 1979).

1.2 GAMETOGENESIS

1.2.1 Fathead minnows may spawn in as little as four to five months after hatching. Little is known about gonadal development during this time, but oogonia are identifiable in ovaries of very young juveniles and oocytes are present in older juveniles (Grizzle 1979). However, since fathead minnows are fractional spawners, all oocyte developmental stages are seen in a mature ovary: (1) oogonia, (2) primary growth stage oocytes, (3) cortical alveolus stage oocytes, and (4) early and late stage vitellogenic oocytes.

1.2.2 Oogonia (12-20 μm diameter)

1.2.2.1 Oogonia (and/or the smallest oocytes present, see Selman *et al.* 1993) are small cells occurring in groups or nests along with similar-sized and larger primary oocytes (Fig. E.1; figures follow Appendix E). Oogonia have a large nucleus with a few variably sized nucleoli and a relatively narrow rim of cytoplasm. They are not surrounded by follicular cells.

1.2.3 Primary growth stage oocytes (primary oocytes, 12-170 μm diameter)

1.2.3.1 Early primary growth oocytes (12-35 μm diameter) are often seen next to other primary oocytes (Fig. E.1). They are still in nests and not completely surrounded by follicular cells, which are also present in the nests. Early primary growth oocytes have a round or oval

nucleus with a few variably-sized nucleoli. The cytoplasm contains no cortical alveoli or yolk bodies.

1.2.3.2 Late primary growth oocytes (35-170 μm diameter) have exited from cell nests and become completely surrounded by squamous follicle cells (Fig. E.2). Both cytoplasmic and nuclear (germinal vesicle) volumes increase considerably during the primary growth phase, as do the numbers of nucleoli, which tend to lie close to the nuclear envelope.

1.2.4 Cortical alvelous stage oocytes (170-425 μm diameter)

1.2.4.1 Cortical alveolus oocytes are characterized by the appearance of cortical alveoli ("yolk vesicles") and, in some species, small lipid droplets, in the cytoplasm. (It has not been clearly established that fathead minnow oocytes have distinct lipid droplets). In early cortical alveolus oocytes, it is possible to observe only one or two cortical alveoli; larger oocytes have many cortical alveoli distributed throughout the cytoplasm (Figs. E.2, 3). The centrally positioned germinal vesicle is oval during this stage and has numerous peripherally located nucleoli of various sizes. The vitelline envelope (zona radiata) is clearly visible in even the smallest cortical alveolus oocytes, and the follicle cells are squamous in early, and cuboidal in late, cortical aveolar oocytes.

1.2.5 Vitellogenic oocytes (425-1070 μm diameter)

1.2.5.1 The initiation of vitellogenesis represents the next oocyte developmental stage which is characterized by the accumulation of eosinophilic yolk bodies in the ooplasm. At first the yolk bodies are much smaller than cortical alveoli and mostly dispersed among them, especially in the perinuclear cytoplasm (Fig. E.3). As the oocyte grows, the yolk bodies become larger and more numerous and displace the cortical alveoli, pushing them to the periphery of the oocyte (Figs. E.3, 6). In the late vitellogenic oocyte (800-1070 μm diameter) the germinal vesicle also appears to move toward the periphery of the oocyte and then disappear entirely when the oocyte approaches maturity (Fig. E.3). In vitellogenic oocytes, the vitelline envelope thickens and becomes striated due to the great numbers of pore channels that penetrate through it. The follicle cells are cuboidal with a large nucleus and a prominent round nucleolus. External to the follicular layer lies a thin basal lamina and a theca consisting of squamous thecal cells, capillaries, and a thin connective tissue stroma. In the most mature oocytes observed in tissue sections the yolk bodies coalesce and may become larger than the cortical alveoli but remain numerous rather than joining into a single yolk mass.

1.3 STAGING OF OVARIES

1.3.1 Staging of ovaries of fractional spawners, such as the fathead minnow, has generally been based upon classification of the most mature oocytes present in the histological section (Selman and Wallace 1986; Leino *et al.* 1990; Selman *et al.* 1993; Leino and McCormick

1997; Shimizu 1997; Jensen *et al.* 2001). Such a classification, based on the studies above, is presented in Table E.1. In fathead minnows, stage 5 oocytes are rarely observed unless females are sampled while spawning. In other words, late stage 4 oocytes are apparently only hours away from being ovulated.

Table E.1. Histological stages of fathead minnow ovarian development.

Stage	Characteristics
1. Primary growth	Oogonia and primary oocytes
	1a Oocytes in nests; small cytoplasmic volume (Fig. E.1)
	1b Oocytes larger, out of nests, surrounded by follicle cells; many pleiomorphic nucleoli bordering the nuclear envelope (Fig. E.2)
2. Cortical alveolus	Appearance of cortical alveoli and, possibly, small lipid droplets (Fig. E.3)
3. Early vitellogenic	Appearance of yolk bodies: initially few and small; ultimately many and variably-sized; centrally located germinal vesicle is round to oval with several peripheral nucleoli (Fig. E.3)
4. Late vitellogenic	Germinal vesicle loses nucleoli, moves towards the periphery and breaks down; yolk bodies frequently fill the entire center of the oocyte and a germinal vesicle may not be evident (Fig. E.3)
5. Mature/spawning oocyte	Germinal vesicle breakdown complete; yolk bodies fuse and may become larger than cortical alveoli

1.4 OVARIAN STAGES IN NORMAL, REPRODUCING FEMALES

1.4.1 Under the test conditions described herein (25 ± 1° C, 16:8-h light:dark photoperiod) most female fathead minnows spawn every 3 to 4 d (Jensen *et al.* 2001). The following description considers variations in ovarian histology during a typical 3 d spawning cycle.

1.4.2 Day 0 post-spawn

1.4.2.1 Ovaries sampled within about 8 h after spawning (which usually takes place in the early morning) have returned to late stage 3. Postovulatory follicles (corpora lutea) are

numerous. In some ovaries most of these follicles are collapsed while in others the follicle lumen is open (Figs. E.4, 5). In goldfish, postovulatory follicles have lumina for at least 10 h after ovulation, but the follicles collapse by 30 h post-ovulation (Nagahama *et al.* 1976). This phenomenon may occur more rapidly in fathead minnows.

1.4.3 Day 1 post-spawn

1.4.3.1 By 1 d post-spawn, ovaries have progressed to stage 4. The post-ovulatory follicles tend to be smaller and have thinner walls than at day 0. Some are vacuolated and appear to be breaking down (Fig. E.6).

1.4.4 Day 2 post-spawn

1.4.4.1 By 2 d post-spawn, ovaries tend to be at late stage 4. Postovulatory follicles are often difficult to identify. Those that are present are small and highly vacuolated.

1.4.5 Day 3 post-spawn

1.4.5.1 By 3 d post-spawn, ovaries are at stage 5 and appear spawning ready.

1.4.6 Atretic follicles

1.4.6.1 If an EDC affects ovarian development and/or spawning, a logical histological feature to check for is an increase in numbers of atretic follicles (preovulatory atretic follicles, POAFs). Follicular atresia in experimental populations of reproducing fathead minnows is generally at a low level (Jensen *et al.* 2001). McCormick *et al.* (1989) reported a mean atresia level (in controls) of 1.6% (range 0-11.6%, n=10) in their experiments, and Miles-Richardson *et al.* (1999b) reported a level of 4.6% (range 0-12%, n=7). Most females have a very low incidence of POAFs, although a small number of control specimens may have higher levels. For example, in the McCormick *et al.* (1989) study, nine ovaries had atresia levels of 0-1.3% and one ovary had a substantially higher level of 11.6%. In a recent study, of 27 "control" ovaries examined, only three had relatively high levels of atresia (MED, unpublished data). Examination of these ovaries showed them to be otherwise histologically normal, at prespawning late stage 4, with atresia mostly or entirely of the most mature follicles (Fig. E.7). Overall, follicular atresia in the 10-12% range may be part of a normal process in some fathead minnows some time during longer spawning periods.

2. NORMAL GONADAL HISTOLOGY IN REPRODUCTIVELY- MATURE MALES

2.1 GENERAL STRUCTURE

2.1.1 The testes are a pair of elongated white organs situated in the dorsal body cavity. Like the ovaries they are suspended by a peritoneal mesentary. Peritoneum covers the testes and

consists of a layer of simple squamous epithelium and a thin connective tissue capsule, the tunica albuginea (Grizzle 1979). Connective tissue septa that separate the seminiferous tubules are continuous with the tunica albuginea. The seminiferous tubules contain the germinal epithelium that ultimately gives rise to spermatozoa (Fig. E.8). In mature testes spermatozoa are present in the lumina of seminiferous tubules and of ampullae that are similar to seminiferous tubules, but lack germinal epithilium (Grizzle 1979). Ampullae empty into the ductus deferens.

2.2 GAMETOGENESIS

2.2.1 The germinal epithelium of fathead minnows has an apparently random distribution of spermatogonia along the entire length of the tubule, the so-called "unrestricted" type of testis (Grier 1981). Spermatogonia are located in small peripheral cysts in the tubule; these cysts enlarge and extend toward the tubule lumen as spermatogenesis proceeds. Five stages of germ cell development are readily identified in the fathead minnow: (1) primary spermatogonia, (2) secondary spermatogonia, (3) primary spermatocytes, (4) secondary spermatocytes, and (5) spermatids and spermatozoa (Smith 1978; Grizzle 1979; Jensen *et al.* 2001 and Fig. E.8).

2.2.2 In addition to the germ cells, two other principal cell-types are present in the testes: Sertoli cells and interstitial cells of Leydig. Sertoli cell bodies, the part of the cell that contains the nucleus, are small and difficult to locate. At high magnification these cell bodies are often triangular-shaped structures situated near the outer rim of the seminiferous tubule (Fig. E.9). The elongate euchromatic nucleus often exhibits a single nucleolus. Processes of a Sertoli cell envelope a cluster of developing germ cells derived from a single primary spermatogonium to form a cyst. These cytoplasmic processes are not usually visible with the light microscope (Grizzle 1979). Numerous polyhedral-shaped Leydig cells are found, usually in groups, in connective tissue spaces between seminiferous tubules (Fig. E.9). They typically have an oval heterochromatic nucleus and a narrow rim of cytoplasm.

2.3 STAGING OF TESTES

2.3.1 Staging of the testes of fathead minnows has been based on the degree of germ cell differentiation (Ankley *et al.* 2001). The presence or absence of certain stages in a histological section, then, can be used to judge the state of testicular maturity. However, it may be advantageous to also consider the relative size and sperm content of the seminiferous tubules (Smith 1978; Leino *et al.* 1990; Gimeno *et al.* 1998) as in Table E.2.

Table E.2. Histological stages of fathead minnow testicular development.

Stage	Characteristics
1. Resting germ cells	No development
2. Spermatogonia	2a Primary spermatogonia: Large cells near edges of tubule; have a lightly staining nucleus with a prominent nucleolus 2b Secondary spermatogonia: Clusters of medium-sized cells with a round, lightly basophilic nucleus; cluster or cyst is the result of several mitotic divisions of a primary spermatocyte.
3. Spermatocytes	3a Primary spermatocytes: Smaller cells with smaller, more basophilic nuclei than spermatogonia: will undergo meiosis I to produce secondary spermatocytes. 3b Secondary spermatocytes: Small cells with smaller, more basophilic nuclei than primary spermatocytes: will undergo meiosis II to produce spermatids.
4. Spermatids and some spermatozoa in lumen of seminiferous tubule; small tubule lumen	Spermatids have a small, intensely basophilic nucleus; they mature into spermatozoa
5. Abundant sperm in an expanded lumen	Figs. E.10, 11

2.4 TESTICULAR STAGES IN NORMAL, REPRODUCING MALES

2.4.1 During a typical 3 to 4 d spawning cycle the testes do not seem to regress to an earlier stage as ovaries do. Examination of testes at 0, 1, 2, and 3 d after a spawning event indicate that, just after spawning, certain seminiferous tubules or regions of these tubules become largely depleted of sperm and have a thin germinal epithelium. Other tubules, however, have a thick germinal epithelium or abundant sperm, or both. Seemingly, sperm production is unlikely to be diminished during normal laboratory spawning.

3. ROUTINE METHODS FOR EVALUATING EDC-INDUCED HISTOLOGICAL CHANGES IN FATHEAD MINNOW GONADS

3.1 OVARIES

3.1.1. Gonadal staging, i.e., as described for ovaries in Table E.1, is a fundamental method for revealing major effects of EDCs on these organs. For example, if ovaries of control fish are stage 4, and those of EDC-exposed fish are stage 3, the EDC has produced a major effect on ovarian development. Gonadal staging is recommended as the first step in the histological evaluation of EDC effects.

3.1.2 Other methods used to describe and evaluate EDC effects on gonads are still evolving based on the kinds of histopathological changes that are being observed, many for the first time. Perhaps the most important of the published methods to attempt to quantify a histological change, is counting the numbers of follicles in various stages of development (Smith 1978; Miles-Richardson *et al.* 1999a,b; Jensen *et al.* 2001). Counting and staging individual follicles provides information on the percentage of particular stages present. For example, treatment of fathead minnows with 10 nM of β-estradiol for 14 d resulted in a greatly increased percentage of primary follicles and a decreased percentage of mature follicles (Miles-Richardson *et al.* 1999a). A recommended method involves counting 100 follicles from sections taken from between 500 µm into the ovary and its midline, and calculating the percentage of each follicular stage present. A similar method has also been employed to assess the severity of oocyte atresia (McCormick *et al.* 1989). This investigation determined that a critical mean percentage of 20% atretic follicles affected spawning success in groups of fathead minnows exposed to acidified water. Note that this percentage is greater than the 10-12% maximal atresia occasionally seen in normally spawning females.

3.2 TESTES

3.2.1 As is the case for ovaries, testicular staging (Table E.2) represents the initial step in evaluating the histological effects of EDCs. Also, as with ovaries, testicular staging will likely reveal only EDC effects that profoundly influence testicular maturation.

3.2.2 Certain quantitative methods have been employed to describe more subtle changes in testicular histology. The first involves an assessment of the percentage of each testicular stage present, such as primary and secondary spermatogonia and spermatocytes. This information can be used to determine whether any of the stages has an atypical distribution. Unlike with ovaries, the relatively small and more numerous testicular germ cells are difficult to count properly without an ocular grid or similar device. Smith (1978) employed an ocular grid to evaluate testicular developmental stages, counting 100 cells in each of three sections per fish. It is important to include different regions from the same testes because testicular histology sometimes varies from one area to another, compared to ovaries which seem to have a rather uniform distribution of oocyte stages throughout.

3.2.3 The second quantitative method involves measurement of the tubule diameter. Certain EDCs may enhance or decrease sperm production. Histologically this may manifest itself as a enlargement or reduction in the mean diameter of seminiferous tubules. Smith (1978) and Gimeno *et al.* (1998) described methods to measure and quantify changes in tubule diameters and relate these changes to sperm production. Tubule diameters should be measured in several testicular regions, for the reasons mentioned above.

3.2.4 Special considerations for testes

3.2.4.1 Some EDC-induced histopathological changes in testes are difficult to study by light microscopy because of the small sizes of the affected cells. This is particularly evident with Sertoli cells which may undergo major EDC-induced changes in morphology that are difficult to see with the light microscope (Miles-Richardson *et al.* 1999a,b). These investigators used electron microscopy to describe changes in Sertoli cells in conjunction with spermatocyte necrosis. Spermatocyte necrosis appears to be a common result of EDC exposure (Miles-Richardson *et al.* 1999a,b; Ankley *et al.* 2000; Länge *et al.* 2001) and electron microscopy is of distinct benefit in supplementing light microscope-based descriptions of this pathology. Electron microscopy may also be of benefit when studying the effects of EDCs on another small testicular cell, the Leydig cell. Although it is reasonable to assume that EDCs may affect Leydig cells, no studies have yet been published on this subject in fish.

3.2.4.2 Major changes in the testes, occasionally even making them difficult to identify in histological sections, can be produced by EDCs. For example, Länge *et al.* (2001) reported ova-testes and frank testicular atrophy in fathead minnows after long-term exposures to ethinylestradiol. More work needs to be done on describing the major histopathological events leading to testicular atrophy.

4. REFERENCES

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.

Gimeno S, Komen H, Jobling S, Sumpter JP, Bowmer T. 1998. Demasculinisation of sexually mature male common carp, *Cyprinus carpio*, exposed to 4-*tert*-pentylphenol during spermatogenesis. *Aquat Toxicol* 43:93-109.

Grier HJ. 1981. Cellular organization of the testis and spermatogenesis in fish. *Amer Zool* 21:345-357.

Grizzle J. 1979. Anatomy and histology of the golden shiner and fathead minnow. PB-294-219. U.S. Dept. of Commerce, National Technical Information Service, Alabama Agricultural Experiment Station, Auburn, AL, USA, 163 pp.

Jensen KM, Korte JJ, Kahl MD, Pasha MS, Ankley GT. 2001. Aspects of basic reproductive biology and endocrinology in the fathead minnow (*Pimephales promelas*). *Comp Biochem Physiol* C 128:127-141.

Länge R, Hutchinson TH, Croudace CP, Siegmund F, Schweinfurth H, Hampe P, Panter GH, Sumpter JP. 2001. Effects of the synthetic estrogen 17α-ethinylestradiol on the life-cycle of the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 20:1216-1227.

Leino RL, McCormick JH, Jensen KM. 1990. Multiple effects of acid and aluminum on brood stock and progeny of fathead minnows, with emphasis on histopathology. *Can J Zool* 68:234-244.

Leino RL, McCormick JH. 1997. Reproductive characteristics of the ruffe, *Gymnocephalus cernuus*, in the St. Louis River estuary on western Lake Superior: a histological examination of the ovaries over one annual cycle. *Can J Fish Aquat Sci* 54:256-263.

McCormick JM, Stokes GN, Hermanutz. 1989. Oocyte atresia and reproductive success in fathead minnows (*Pimephales promelas*) exposed to acidified hardwater environments. *Arch Environ Contam Toxicol* 18:207-214.

Miles-Richardson SR, Kramer VJ, Fitzgerald SD, Render JA, Yamini B, Barbee SJ, Giesy JP. 1999a. Effects of waterborne exposure of 17β-estradiol on secondary sex characteristics and gonads of fathead minnows (*Pimephales promelas*). *Aquat Toxicol* 47:129-145.

Miles-Richardson SR, Pierens SL, Nichols KM, Kramer VJ, Snyder EM, Snyder SA, Render JA, Fitzgerald SD, Giesy JP. 1999b. Effects of waterborne exposure to 4-nonylphenol and nonylphenol ethoxylate on secondary sex characteristics and gonads of fathead minnows (*Pimephales promelas*). *Environ Res* 80:S122-S137.

Nagahama Y, Chan K, Hoar WS. 1976. Histochemistry and ultrastructure of pre- and post-ovulatory follicles in the ovary of the goldfish, *Carassius auratus*. *Can J Zool* 54:1128-1139.

Selman K, Wallace RA. 1986. Gametogenesis in *Fundulus heteroclitus*. *Amer Zool* 26:173-192.

Selman K, Wallace RA, Sarka A, Qi X. 1993. Stages of oocyte development in the zebrafish, *Brachydanio rerio*. *J Morphol* 218:203-224.

Shimizu A. 1997. Reproductive cycles in a reared strain of the mummichog, a daily spawner. *J Fish Biol* 51:724-737.

Smith, RJF. 1978. Seasonal changes in the histology of the gonads and dorsal skin of the fathead minnow, *Pimephales promelas*. *Can J Zool* 56:2103-2109.

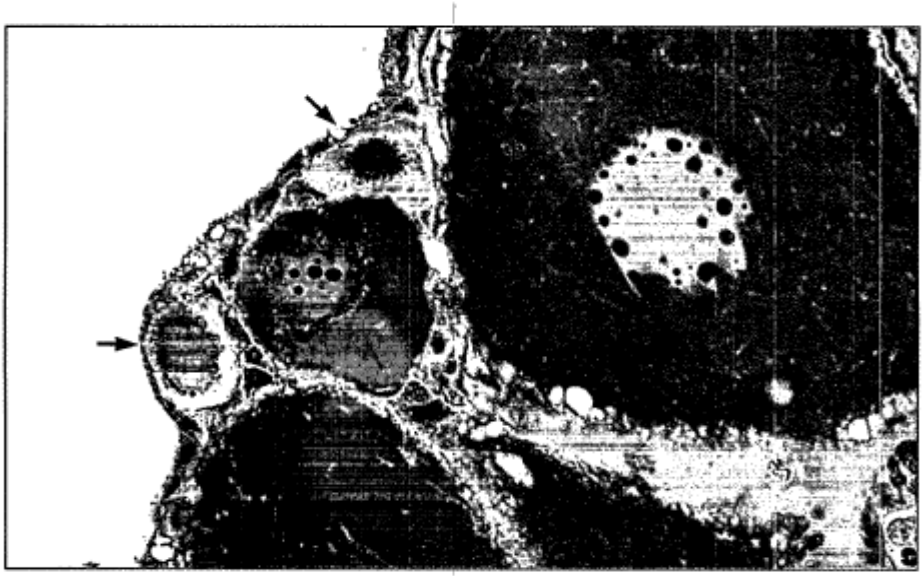

Figure E.1. Oogonia or small primary oocytes (arrows) in nest with larger primary oocytes. (1170x)

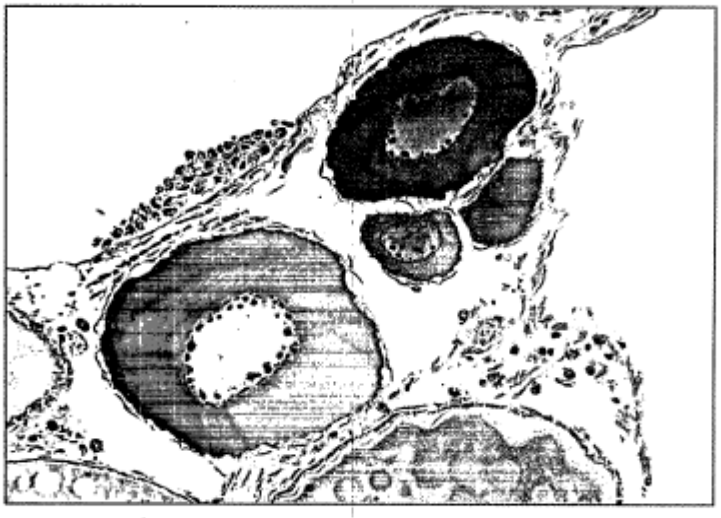

Figure E.2. Late primary growth oocytes of various sizes. Late primary oocytes are surrounded by squamous follicle cells. (380x)

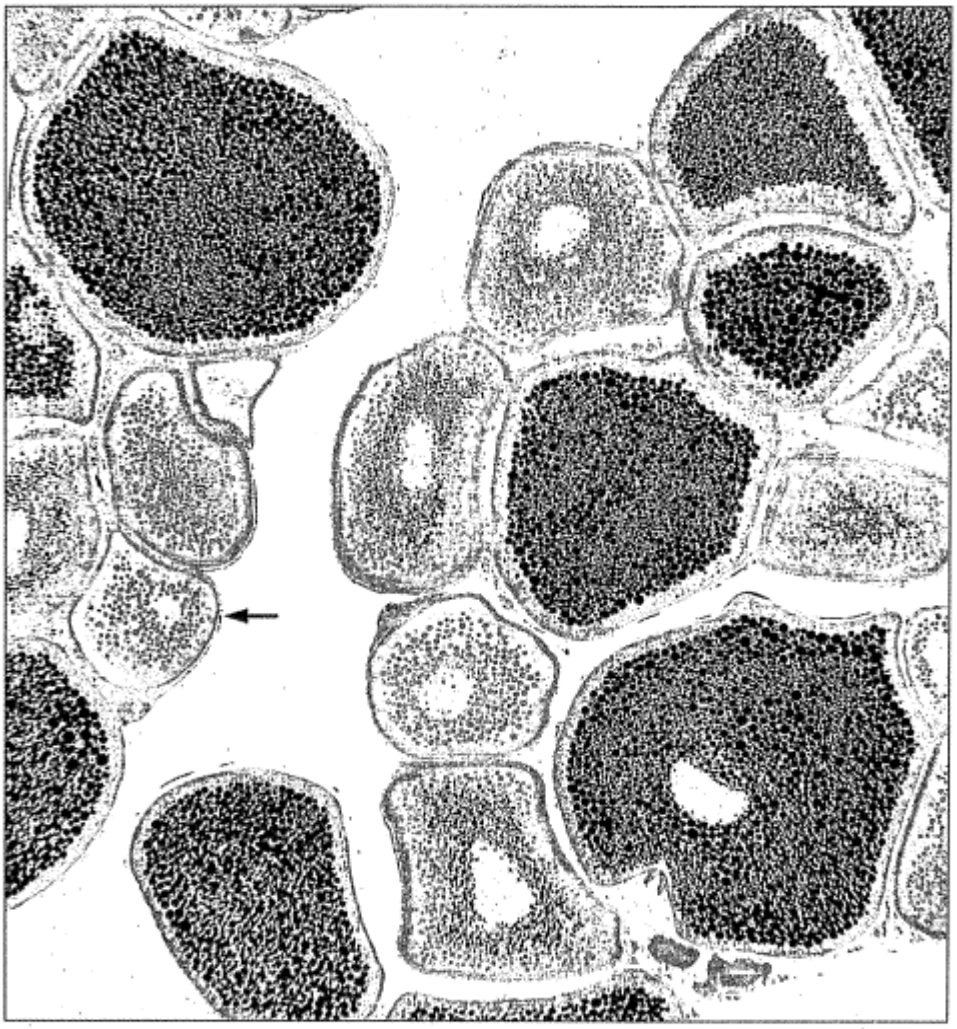

Figure E.3. Part of ovary containing early and late vitellogenic oocytes with densely-stained yolk bodies. Note single cortical alveolus stage oocyte (arrow). (65x)

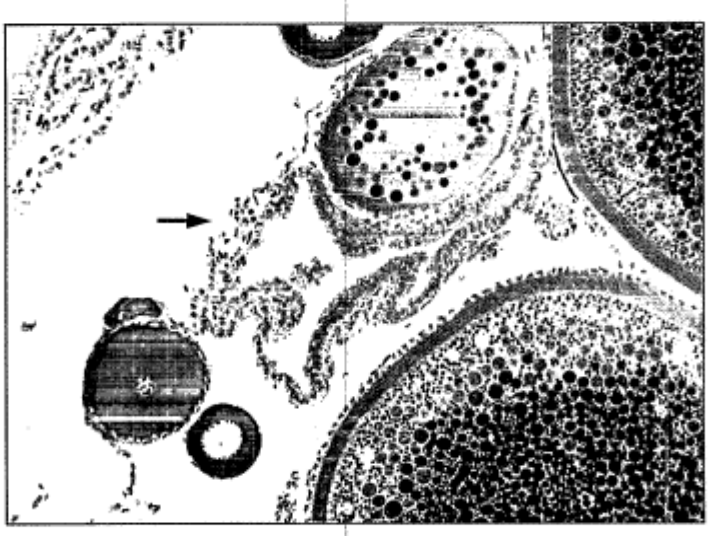

Figure E.4. Post-ovulatory follicle (corpus luteum) with an open lumen from a day 0 post-spawn ovary (see text). (190x)

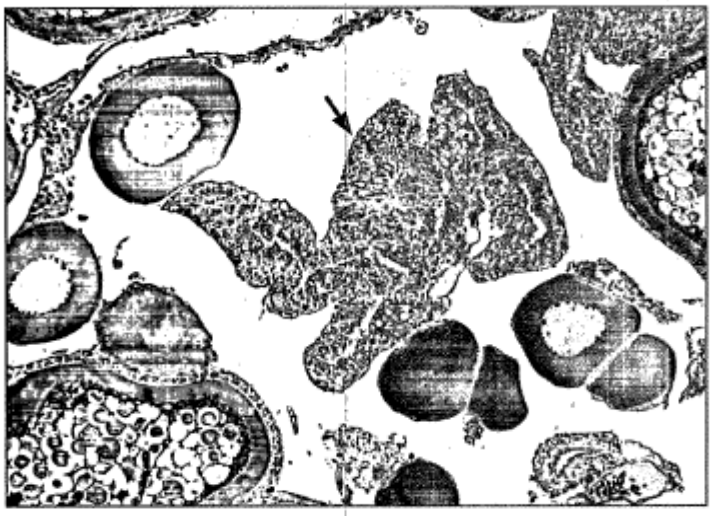

Figure E.5. Collapsed corpus luteum from a day 0 post-spawn ovary. (190x)

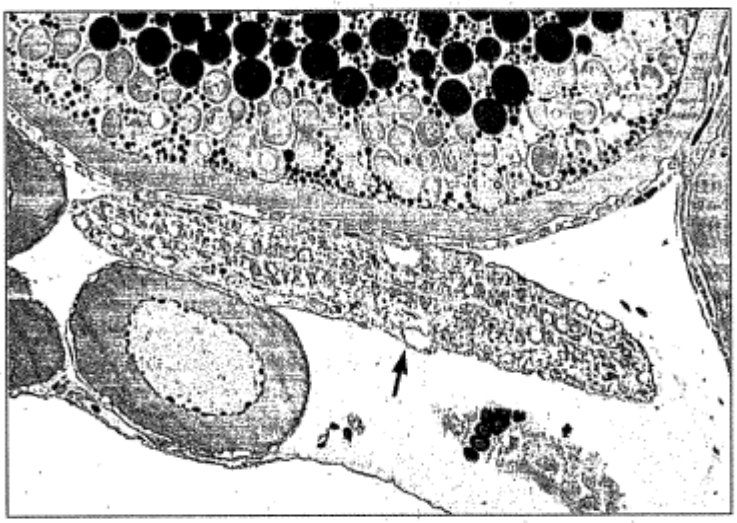

Figure E.6. Vacuolated corpus luteum from a day 1 post -spawn ovary. (360x)

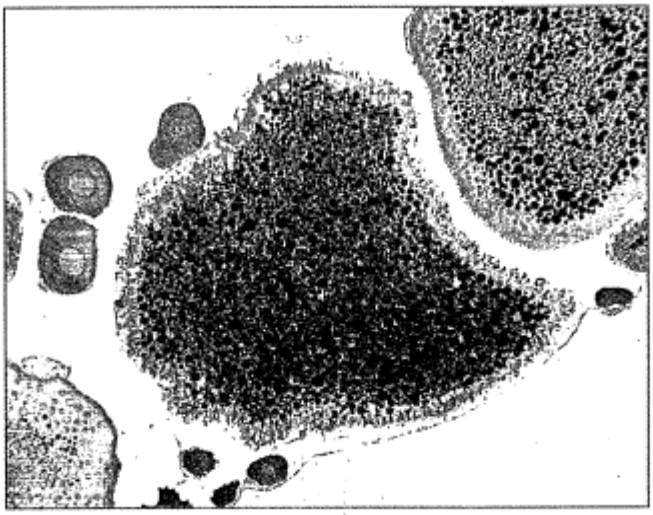

Figure E.7. Pre-ovulatory atretic follicle. (90x)

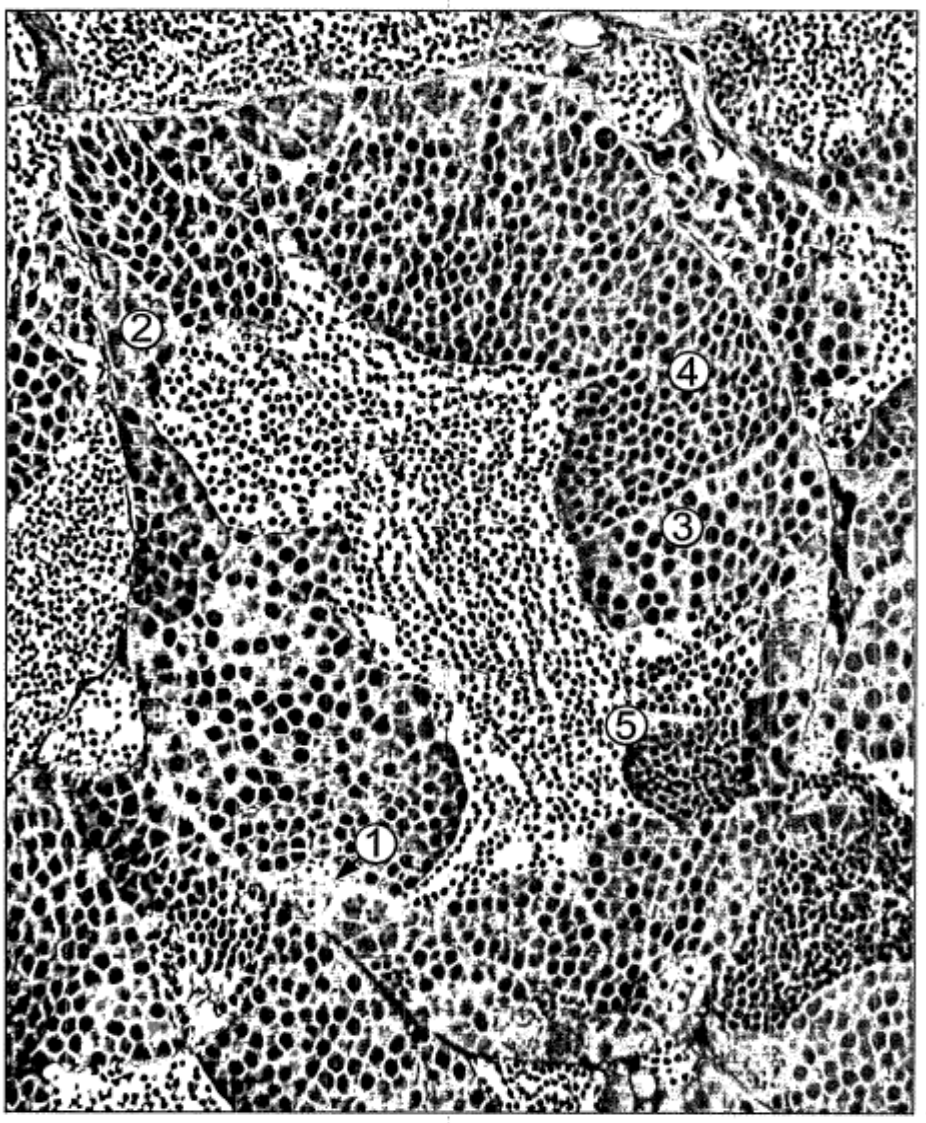

Figure E.8. Section of a seminiferous tubule showing various developmental stages: (1) primary spermatogonia, (2) secondary spermatogonia, (3) primary spermatocyte, (4) secondary spermatocyte, (5) spermatids, late cyst above and early cyst beneath. (645x)

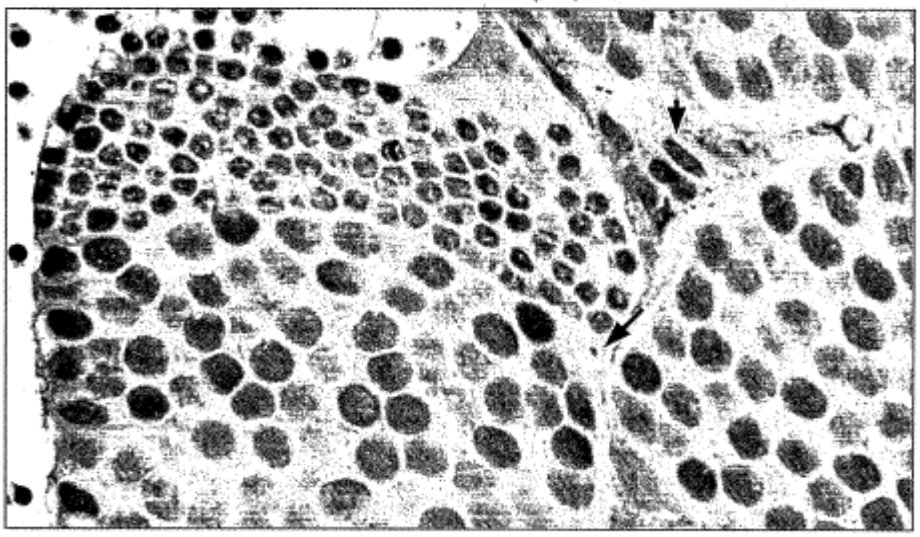

Figure E.9. Sertoli cell (long arrow) and interstitial cells of Leydig (short arrow) in testis. (1680x)

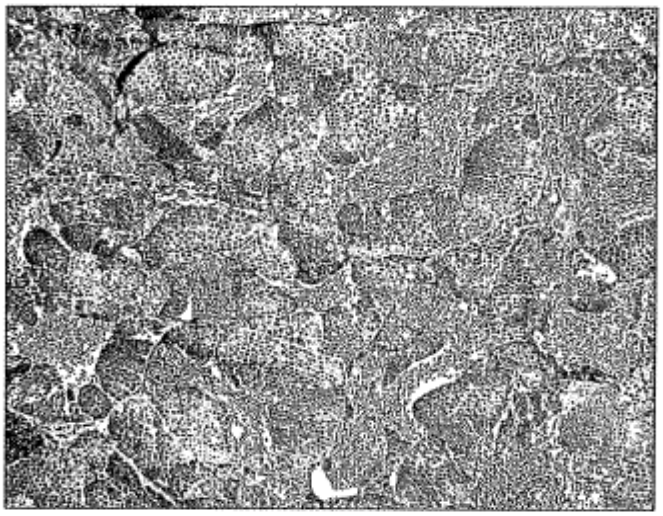

Figure E.10. Stage 5 testis with thick germinal epithelium and sperm-filled lumina. (160x)

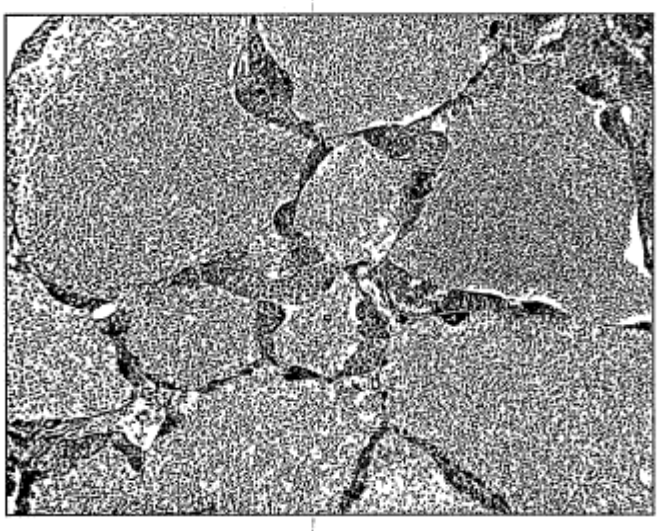

Figure E.11. Stage 5 testis with thin germinal epithelium and expanded lumina. (165x)

United States
Environmental Protection
Agency

Please make all necessary changes on the below label, detach or copy, and return to the address in the upper left-hand corner.

If you do not wish to receive these reports CHECK HERE ☐; detach, or copy this cover, and return to the address in the upper left-hand corner.

PRESORTED STANDARD
POSTAGE & FEES PAID
EPA
PERMIT No. G-35

National Health and Environmental
Effects Research Laboratory
Duluth, MN 55804

Official Business
Penalty for Private Use
$300

EPA/600/R-01/067
June 2002