Application of an Integrated Testing Strategy to the U.S. EPA Endocrine Disruptor Screening Program

Catherine E. Willett,*1 Patricia L. Bishop,* and Kristie M. Sullivan†

*Regulatory Testing Division, People for the Ethical Treatment of Animals, Norfolk, Virginia 23510; and †Toxicology and Regulatory Testing, Physicians Committee for Responsible Medicine, Washington, DC 20016

1To whom correspondence should be addressed at People for the Ethical Treatment of Animals, Regulatory Testing Division, 501 Front Street, Norfolk, VA 23510.

Fax: (617) 522-3487. E-mail: katew@peta.org.

Received February 12, 2011; accepted May 23, 2011

Key Words: endocrine testing; alternative methods; integrated testing strategy; testing framework.

New approaches to generating and evaluating toxicity data for chemicals are needed to cope with the ever-increasing demands of new programs. One such approach involves the use of an integrated testing and evaluation strategy based on the specific properties and activities of a chemical. Such an integrated strategy, whether applied to existing or future programs, can promote efficient use of resources and save animals. We demonstrate the utility of such a strategy by applying it to the current U.S. Environmental Protection Agency Endocrine Disruptor Screening Program (EDSP). Launched in October 2009, the EDSP utilizes a two-tiered approach, whereby each tier requires a battery of animal-intensive and expensive tests. Tier 1 consists of five in vitro and six in vivo assays that are intended to determine a chemical’s potential to interact with the estrogen (E), androgen (A), or thyroid (T) hormone pathways. Tier 2 is proposed to consist of multigenerational reproductive and developmental toxicity tests in several species and is intended to determine whether a chemical can cause adverse effects resulting from E, A, or T modulation. In contrast to the existing EDSP structure, we show, using the pesticide atrazine as an example, that a multilevel testing framework combined with an integrated evaluation process would significantly increase efficiency by minimizing testing.

Key Words: endocrine testing; alternative methods; integrated testing strategy; testing framework.

In response to a 1996 amendment to the Federal Food, Drug, and Cosmetic Act, the Environmental Protection Agency (EPA) has developed the Endocrine Disruptor Screening Program (EDSP) to evaluate chemicals for possible effects on the estrogen (E), androgen (A), and thyroid (T) systems in humans and wildlife. The current EDSP is organized into two tiers. Tier 1 consists of five in vitro and six in vivo assays that are intended to “identify substances that have the potential to interact with the EAT (estrogen/androgen/thyroid) hormonal systems . . .,” (EPA, 2009) (Table 1). The putative Tier 2 battery consists of developmental and reproductive toxicity tests in several vertebrate species (Table 2) and is designed to identify and establish dose-response relationships for any adverse endocrine-related effects. The EPA has stated that it intends to use a weight-of-evidence (WoE) approach to evaluate Tier 1 results (EPA, 2010a; 74 FR 17560) and determine which, if any, of the Tier 2 tests are necessary.

Conducting all of the 11 EDSP Tier 1 tests for one chemical would require a minimum of 520 animals (Table 1) (Willett and Sullivan 2009) and cost between $335,100 and $964,250 (Organization for Economic Co-operation and Development [OECD], 2010). Phase I of the program required testing of 67 chemicals, including 58 pesticide active ingredients and 9 High Production Volume chemicals used as pesticide inert ingredients, most of which already have a wealth of data associated with them. During the registration process, pesticides undergo extensive testing, including reproductive and chronic/life cycle studies in rodents, fish, and birds, as well as metabolism and pharmacokinetics studies. Together, these tests kill thousands of animals and include many of the endpoints addressed in the presumptive EDSP Tier 2 tests (e.g., reproductive toxicity in one or more generations).

With the vast number of substances slated for evaluation in subsequent phases of the EDSP, adoption of a more streamlined testing process is essential. Efficiency of testing can be increased by consideration of existing information, prioritization according to defined criteria, and tiered testing with clear off-ramps based on information goals. Although applying these principles to endocrine-related modes of action can be difficult due in part to the diverse and complex modalities that may be involved, a testing framework can be constructed to guide evaluation of available information and further testing. Following such a framework, combined with iterative integrated assessment at each step, can save resources, time and animals, increase consistency in decision making, and ensure efforts are directed toward substances with the most
potential to affect the hormonal pathways of humans or wildlife.

INTEGRATED TESTING STRATEGY FOR DECISION MAKING

Rather than a default application of the entire battery of Tier 1 assays, we propose a multilevel testing framework combined with iterative assessment at each step (Fig. 1). The testing framework is similar to the structure of the OECD Conceptual Framework (CF) for the Testing and Assessment of Endocrine Disrupting Chemicals (Table 3). Although the CF is useful in organizing available information and identifying potential assays that may be used to gather new information, it is not designed to be a linear testing strategy and does not provide for interim WoE decision making. Like the OECD CF, our strategy assigns information to different levels but additionally provides an opportunity for assessment and decision making at each level with regard to whether to continue, and if so, what further testing is most relevant. This process of sequential data interpretation and assessment relies on a WoE approach that considers both mechanistic and apical information; the more data that support a particular conclusion, the more reliable that conclusion will be. This tailored iterative process thus forms the basis of our integrated testing strategy (ITS).

Level 1: Sorting and Prioritization

Initial prioritization of a chemical for testing can be based on its physical/chemical properties (e.g., solubility, partitioning
coefficient); production volumes; presence and fate in the environment; the likelihood of exposure to humans and wildlife; and the life stages at which those exposures may occur. As an example, the draft list of 134 chemicals EPA identified for testing in Phase 2 of the EDSP (75 FR 70248) includes sulfosate, a pesticide for which the registration was canceled and production ceased in 2002. According to our strategy, a chemical such as this, one that is no longer produced, is not persistent in the environment and was estimated to be present in drinking water at concentrations well below action levels for both acute and chronic exposure when it was being actively used (66 FR 48601), would receive a low priority for testing.

Toxicokinetic data (e.g., bioavailability, half-lives for absorption and elimination), read across from in vivo results obtained with chemically related substances, and (quantitative) Structure Activity Relationship [(Q)SAR] models/expert systems are also valuable for preliminarily assessing a chemical’s potential to
TABLE 3

The OECD Conceptual Framework for Endocrine Disruptor Screening (2002 updated)

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1:</td>
<td>sorting and prioritization based on existing information</td>
<td>Physical and chemical properties, Human and environmental exposure, Hazard (available toxicological data), QSAR for metabolism, bioaccumulation, etc.</td>
</tr>
<tr>
<td>Level 2:</td>
<td>in vitro assays providing mechanistic data</td>
<td>Estrogen, androgen, and thyroid receptor binding, Transcriptional activation via ER (TG 455), AR, TR, Aromatase, Steroidogenesis, QSAR for ER or AR activity, High-throughput screens, Thyroid function, Fish hepatocyte vitellogenin</td>
</tr>
<tr>
<td>Level 3:</td>
<td>in vivo assays providing data about single mechanisms and effects</td>
<td>Uterotrophic (TG 440), Hershberger (TG441), Amphibian metamorphosis assay (TG 231)*, Xenopus embryo thyroid signaling assay (Fish vitellogenin—e.g., TG 230)</td>
</tr>
<tr>
<td>Level 4:</td>
<td>in vivo assays providing data about multiple mechanisms and effects</td>
<td>Enhanced TG 407, Male and female pubertal assays, Adult intact male, Fish short-term reproductive assay (TG 229), Fish sexual development test (draft TG)</td>
</tr>
<tr>
<td>Level 5:</td>
<td>in vivo assays providing data on effects from endocrine and other mechanisms</td>
<td>One- and two-generation reproductive toxicity in rodents (TG 415, 416), extended one-generation reproductive toxicity study (draft TG), Partial and full life cycle assays in fish, birds, amphibians (e.g., TG 206)</td>
</tr>
</tbody>
</table>

*Although we have included the amphibian metamorphosis assay in Level 3 since a positive result is primarily indicative of a thyroid mode of action, it is arguable that this assay is more appropriate in Level 4.

Therefore, (Q)SARs are relevant in both Level 1 for priority setting and in Level 2 as predictive models for mechanism.

In addition to physicochemical information, existing information regarding known toxicities or mechanisms or modes of action (MoA) can be useful in prioritizing chemicals. (In this paper, MoA refers to mode of action— the biological pathway [e.g., estrogenic, androgenic], versus the more specific mechanism of action [e.g., estrogen receptor binding].) For example, if a chemical is a known mutagen, carcinogen or neurotoxin or another MoA has been well characterized and no-effect or lowest-effect levels determined for that toxicity, additional testing for endocrine activity is of little regulatory value and therefore that chemical would be assigned a low priority.

Level 1 information may not always help to characterize a substance’s potential for endocrine activity and will not likely be a final determinant as to whether a substance should undergo further scrutiny. However, collection of this information is essential both to prioritize substances for testing and to guide the investigator in constructing a testing strategy for a particular substance or group of substances.

Level 2: In Vitro Mechanistic Screens/Tests

Guided by determinations made in Level 1, use of mechanistic (Q)SARs and expert systems—assays such as ER, AR, and TR binding; ER-, AR-, and TR-dependent transcriptional activation; aromatase; steroidogenesis; and fish hepatocyte vitellogenin—can rapidly assess chemicals qualitatively for primary activity and rank them according to relative activity. Assays at this level can also be used to identify possible MoAs, further prioritize substances, indicate potential hazard, and direct possible further testing. An example of using mechanism-based assays to rank chemicals according to potential ED activity was the system developed by Reif et al. (2010), which combined results of 90 in vitro endocrine-related assays from EPA’s ToxCast Program with metabolic activity, bioavailability, and pathway perturbation “scores” to evaluate over 200 substances.

Several (Q)SARs and expert systems exist for estimating ER and AR binding, agonist and antagonist activity, and more complicated reproductive endpoints (reviewed in Cronin and Worth, 2008; Novic and Vracko, 2010). An expert system for predicting ER-binding activity of inert ingredients and antimicrobial pesticides was recently positively reviewed by both OECD and EPA (OECD, 2009). A combination of (Q)SARs, each predicting different reproductive endpoints, was used to identify 5240 of 57,014 chemicals in the European Inventory of Existing Chemical Substances as putative reproductive toxicants (Jensen et al., 2008). These chemicals were then further evaluated by (Q)SAR models predicting ER-binding, ER agonist, and AR antagonist activity to conclude that only 3–5% of the identified reproductive toxicity was due to ER or AR mechanisms. A strong caveat of (Q)SAR use is that each (Q)SAR has a defined applicability domain determined by the

interact with the endocrine system. (Q)SAR models are increasingly viewed as one of the most cost-effective ways to estimate ecological and health effects of chemicals and are particularly helpful when little measured data are available for a chemical. They can be employed for assessment of factors, such as metabolism, fate, and bioconcentration (e.g., Arnot and Gobas, 2003; Cronin et al., 2003; Harju et al., 2007), and for predicting endocrine-disrupting (ED) properties, such as the binding of substances to the estrogen receptor (ER) and androgen receptor (AR) (e.g., Hong et al., 2002; Lill et al., 2005; Schmieder et al., 2003; Walker et al., 2003).
information used to inform it. All (Q)SARs therefore require careful consideration of relevance to a particular chemical.

Similarly, in general, each in vitro assay measures a single mechanism and thus conclusions can be drawn only in the context of what the in vitro assay evaluates. Positive results indicate the possibility of ED effects in vivo via the mechanism tested, but it is important to note that many assays are designed to be overresponsive (biased toward false positives) to minimize the risk that true EDs will go undetected. In the case of negative results at this level, possible ED activity cannot be excluded, as chemicals can affect the endocrine system through mechanisms for which suitable in vitro assays do not yet exist, such as via other receptors (e.g., corticosteroid receptors) or by indirect effects on the hypothalamic-pituitary-gonadal (HPG) axis. In addition, lack of metabolizing systems in most in vitro assays could lead to false negatives. However, if supporting ADME data are available and can rule out the chemical undergoing metabolic activation, then more confidence can be assigned to a negative result at this level. In other cases, ADME information can identify the active metabolite, which can then be used for Level 2 testing. Nevertheless, a lack of activity in multiple tests addressing the same mechanism or pathway gives some confidence that the chemical does not act via that mechanism. Several metabolizing systems are now available for use in conjunction with in vitro assays, which could strengthen confidence in negative results at this level (OECD, 2007). Thus, validation of these systems for regulatory use should be made a priority.

Information from Level 2 can also be used to direct further testing. For example, if a chemical is positive for several tests highlighting a specific pathway at relatively low concentrations and shows little or no activity in assays for other pathways, further testing could be focused on the pathway for which the chemical has been demonstrated to be most active.

**Level 3: In Vivo Mechanistic Tests**

Based on the results of Level 2 testing and existing information, assays such as the uterotrophic, Hershberger, and Xenopus embryo thyroid signaling (Fini et al., 2007) are used when required to confirm any effects observed in the previous level. These tests give a yes/no (qualitative) answer regarding effects and generally include the possibility for some metabolic activation of a chemical. They are designed to be overly sensitive (e.g., use of castrated/immature animals without intact HPG axis) and are usually of short duration.

While these tests typically expose animals to one or few concentrations at a specific point in their life cycle, dosing over a larger range could potentially provide enough information for hazard assessment and would obviate the need for more animal-intensive, higher level in vivo testing.

Positive results associated with a specific ED endpoint would generally indicate a possibility for adverse effects in longer reproductive and developmental studies. Negative results may be sufficient to enable a conclusion of no concern for ED. This will depend upon the WoE and may not be possible in some cases. However, in the presence of negative or positive data from robust Level 4 and 5 assays (often available from previous testing of a substance), further animal testing is not justified. If a thyroid mechanism of action is indicated, it may be more appropriate to consider conducting Level 4 tests, as there are currently no validated in vivo mechanistic assays that are specific to thyroid hormone activity.

**Level 4: In Vivo Tests for Evaluating Multiple Mechanisms and Effects**

Assays at this level are designed to assess endpoints that may be sensitive to more than one mode of ED action and may provide some information on potency, hazard, and risk assessment. They include such animal-intensive studies as the male/female rat pubertal assays, the amphibian metamorphosis, and fish short-term reproduction tests. In a WoE analysis, these Level 4 studies are only conducted if both of the following conditions are met: (1) there are positive results from each of the prior levels or if higher level studies conducted previously indicate effects through endocrine-mediated pathways and (2) additional information is required for regulatory purposes that is not obtainable from the lower level tests. Positive Level 4 results would indicate a likelihood of adverse effects in longer-term studies; however, such studies should only be considered if more information is necessary for hazard assessment in a particular regulatory situation. Negative results at this level can rule out a chemical’s activity.

**Level 5: In Vivo Tests for Evaluating Adverse Outcomes, Dosage Responses, and Risk Assessment**

These animal-intensive studies include mammalian, avian, amphibian, and fish partial or full lifecycle reproductive and developmental toxicity tests. They would be conducted only after (1) lower level testing has been conducted and an endocrine pathway is clearly indicated and (2) dose/response or assessment of a particular apical endpoint is required for hazard or risk assessment.

It is important to note that whenever using existing information in a WoE analysis at any level, evaluation of the quality, relevance, and adequacy of the study data is essential and should be approached systematically (e.g., Klimisch et al., 1997; Schneider et al., 2009).

Unlike our integrated strategy, the current structure of the EDSP does not take into account the type of information evaluated in Level 1, such as physicochemical properties, existing toxicokinetic, or structure information. Additionally, it prioritizes chemicals based solely on exposure potential and, thus, limits the application of an ITS. As the EDSP offers no means for interim decision making, the entire Tier 1 test battery must be performed and results evaluated in a WoE approach before the decision to conduct further testing can be made.
Currently, test order recipients have the option of submitting Other Scientifically Relevant Information (OSRI), which, if deemed by EPA to be directly or functionally equivalent to one or more of the Tier 1 tests, can satisfy the requirement of performing those tests. EPA’s current practice limits consideration of OSRI to whether or not it can substitute for a Tier 1 test; however, OSRI has considerable value when building a WoE conclusion at any level of the iterative decision-making process. Many chemicals, particularly pesticides, have been tested extensively in the past using assays of the Level 4 and 5 types. While older reproductive or developmental toxicity tests, for instance, may not have measured all sensitive ED endpoints now included, this information is useful when combined with lower level results. For example, if a chemical is negative in all mechanistic tests (e.g., those in Levels 2 and 3) and negative in a multigeneration reproduction test, it is highly unlikely that this chemical possesses ED activity. Conversely, determining whether positive reproductive findings are a result of this chemical possesses ED activity. Conversely, determining whether positive reproductive findings are a result of this approach since, in addition to the large volume of toxicological information available for this chemical, atrazine has been tested in assays that are similar or identical to all of the EPA’s Tier 1 tests, making a hypothetical comparison between the two approaches possible. For the purposes of illustration, the following analysis uses existing information for atrazine other than specific endocrine tests until the analysis has reached the level at which that testing would occur.

### EXAMPLE OF ITS FRAMEWORK APPROACH: ATRAZINE

Atrazine was chosen to demonstrate the hypothetical use of this approach since, in addition to the large volume of toxicological information available for this chemical, atrazine has been tested in assays that are similar or identical to all of the EPA’s Tier 1 tests, making a hypothetical comparison between the two approaches possible. For the purposes of illustration, the following analysis uses existing information for atrazine other than specific endocrine tests until the analysis has reached the level at which that testing would occur.

#### Level 1: Prioritization

Atrazine, a chlorotriazine herbicide that has been on the market for over 50 years, is one of the most extensively used agricultural products worldwide (e.g., EPA, 2010b). Every year in the United States, farmers apply more than 70 million pounds to crops before and after planting (Chevrier et al., 2011). Exposure can occur during manufacture, handling and application of the herbicide, and through presence in water and on crops.

Atrazine metabolites are relatively persistent in soil/sediment, with mean aerobic and anaerobic soil half-lives ranging from 58 to 547 days and an aerobic aquatic half-life about twofold higher (Solomon et al., 2008). Even though atrazine is only moderately soluble in water and degrades fairly rapidly, it does not bind strongly to sediments, is moderately to highly mobile in soils, and, therefore, has a high potential for contamination of groundwater and surface waters. Individual drinking water samples collected under the Atrazine Monitoring Program have occasionally exceeded the Maximum Contaminant Level (MCL) for long-term exposure of 3 ppb; in 2009 and 2010, 4 and 7% of the samples, respectively, were above the MCL (EPA, 2011a).

There is a large amount of existing information for atrazine, including extensive investigation of the MoA of its primary toxicity. Atrazine has been shown to affect reproduction in multiple species indirectly via the central nervous system (CNS). Atrazine exposure reduces luteinizing hormone (LH) secretion by directly suppressing gonadotropin-releasing hormone (GnRH) from the hypothalamus (Cooper et al., 2007).

Based on its widespread use, large amounts applied, documented presence and persistence in the environment, and resultant high potential for exposure to humans and wildlife, atrazine would be given a high priority for additional ED testing and evaluation. However, this conclusion is altered by the thorough existing information documenting the herbicide’s adverse effects and primary MoA. Therefore, a Level 1 analysis would indicate that further testing of atrazine is of little regulatory use, and if further information about its potential MoA is required, Level 2 and perhaps Level 3 testing should address those needs.

#### Level 2: In Vitro Mechanistic Testing

**Receptor binding, transcriptional activation, and ToxCast assays.** Atrazine has consistently failed to activate estrogen-dependent reporters in vitro in estrogen-dependent expression systems (Eldridge et al., 2008). Likewise, atrazine failed to stimulate estrogen-dependent cell proliferation using MCF-7 (Fukamachi et al., 2004) or MтТ/Е-2 (Fujimoto, 2003) nor did it induce ER-dependent transcription in T47D (Legler et al., 2002), Chinese hamster ovary (CHO) (Kojima et al., 2004), or yeast cells (O’Connor et al., 2000). Tennant et al. (1994) measured estrogen binding using three different in vitro assays (one similar to the Tier 1 protocol) and reported negative results in all. A recent review of binding data (Cooper et al., 2007) also concluded that atrazine does not bind the ER except at extremely high concentrations.

Atrazine did not bind to the human recombinant AR (Yamasaki et al., 2004). It did not activate AR-dependent transcription in CHO cells (Kojima et al., 2004) and was used in validation of the Tier 1 AR-binding assay (EPA, 2007) as a negative compound, further confirming its inactivity via the AR receptor.

Atrazine was negative in all six ToxCast (ToxCastDB). ER-related assays, in all five ToxCast AR-related assays, and in all four ToxCast TR-related assays (Reif et al., 2010).

**Aromatase and steroidogenesis.** Several in vitro tests have shown that atrazine increases aromatase levels by binding to and inhibiting phosphodiesterase (Roberge et al., 2004; Sanderson et al., 2000, 2001), resulting in elevated cAMP in some human cancer cell lines. Fan et al. (2007) demonstrated that atrazine affects aromatase expression only in cell and tissue types that use the steroidogenic factor-1–dependent ArPII promoter,
which is critically involved in breast cancer oncogenesis. Aromatase activity in human granulosa-lutein cells (which constitutively express aromatase) and in endometrial cells (which do not express aromatase) increased twofold in the former and did not change in the latter, indicating that atrazine is an inducer of aromatase (Holloway et al., 2008).

Atrazine was tested as part of the development and validation of the Tier 1 steroidogenesis (H295R) assay (EPA, 2008; Higley et al., 2010) where it was shown to affect aromatase activity indirectly but not directly, resulting in increased 17-β-estradiol, testosterone, and aromatase activity. It was classified as a “general inducer” of aromatase, qualitatively similar to forskolin but with much weaker activity.

In ToxCast, atrazine was positive in the CellzDirect UGT1A1 assay that measures transcriptional activation of the uridine diphosphate glucuronosyltransferase (UGT) gene in primary human hepatocytes (Reif et al., 2010), supporting the hypothesis that atrazine may affect steroid hormone metabolism.

Based on Level 2 WoE, atrazine does not function via an ER, AR, or TR mechanism of action but may interfere with steroid hormone metabolism. Thus, Level 3 tests, including uterotrophic and Hershberger, fish short-term reproduction, and amphibian metamorphosis tests, are not recommended; Level 4 tests addressing more complex thyroid and multiple mechanisms of action may be more appropriate if additional information is required.

**Level 3: In Vivo Mechanistic Tests**

*Estrogenicity and androgenicity.* In uterotrophic studies using both immature and ovariectomized adult rats, atrazine failed to demonstrate estrogenic activity but did display weak anti-estrogenic activity when co-administered with 17β-estradiol (Tennant et al., 1994; Connor et al., 1996; Yamasaki et al., 2000).

An assessment of atrazine using the castrate version of the Hershberger assay showed no androgenic agonist or antagonist activity of atrazine exposure (Yamasaki et al., 2004). Atrazine showed possible antiandrogenic activity at the highest doses tested (100 mg/kg) in three Hershberger studies cited by EPA in response to OSRI submitted for atrazine (EPA, 2011b).

Atrazine exposure via the diet (up to 1000 ppm or 109 mg/kg/day) or sc injection (up to 10 mg/kg/day) had no significant affect on uterine weight or pituitary LH release in immature female quails (Wilhelms et al., 2006), confirming and expanding earlier studies that showed the absence of estrogen-like effects in the maturing reproductive tracts of male quail administered up to 1000 ppm atrazine (Wilhelms et al., 2005). Although these tests are not considered routine Level 2 tests, these studies address ER and AR mechanisms of action and so the information is considered here.

*Thyroid effects and steroidogenesis.* Some studies reported that atrazine affects sexual development and gonadal differentiation in the frog (*Xenopus laevis*) at low concentrations (Hayes et al., 2002, 2006), while others did not observe significant effects on developing *Xenopus* (Carr et al., 2003; Oka et al., 2008). The result reported by Carr et al. (2003) was confirmed in a larger study conducted concurrently in two laboratories (Kloas et al., 2009). In this study, no effects on any developmental or gonadal parameters were observed when frogs were exposed to atrazine at concentrations of 0.01, 0.1, 1.0, 25, or 100 ppb from day 8 postfertilization until the completion of metamorphosis (a time frame of exposure exceeding that of the Amphibian Metamorphosis Assay required in the Tier 1 battery). In contrast, estradiol administered under similar conditions as a control at a concentration of 0.2 ppb resulted in a significant increase in larvae displaying female or mixed sex gonads compared with untreated controls.

The effect of atrazine on steroid hormone synthesis has been widely studied in vivo. Findings have shown that atrazine decreases circulating testosterone via an indirect effect on aromatase gene expression (Rosenberg et al., 2008; Victor-Costa et al., 2010). A recent study of ex vivo Leydig cells following peripubertal exposure showed that atrazine decreased expression of several genes responsible for steroidogenesis at doses of 50 mg/kg/day and higher, which is likely to be the underlying cause of the decrease in testosterone seen in vivo (Pogrmic et al., 2009).

Level 3 WoE confirms that atrazine does not function via ER, AR, or direct thyroid hormone-related mechanisms and also confirms that atrazine has some effect on steroidogenesis. This retrospective analysis suggests that Level 3 testing provided confirmatory but not new or necessary information regarding potential endocrine activity of atrazine.

**Level 4: In Vivo Mechanistic Testing for Multiple Actions and Effects**

*Pubertal studies in rats.* In a pubertal study carried out by Laws et al. (2000) using female Wistar rats, atrazine delayed vaginal opening (VO) and altered estrous cyclicity at 50 mg/kg/day and above (no-observed-adverse-effect level [NOAEL] of 25 mg/kg/day). Reduced food consumption and body weight did not account for the delay in VO as this effect was not observed in pair-fed controls. Results of the study suggest that atrazine delays onset of puberty by altering hypothalamic-pituitary activity. Suppression of LH secretion caused by direct suppression of GnRH from the hypothalamus was later supported by results of Cooper et al. (2007). Ashby et al. (2002) observed delayed VO and reduced gains in uterine weights in Wistar and Sprague-Dawley (SD) rats; however, this effect was shown to be a delay rather than a complete block in maturation when compared with the effects elicited by the centrally acting GnRH antagonist Antarelix, which was used as a positive control agent.

In male pubertal studies, atrazine caused delayed onset of puberty with decreased sex organ weights at high doses, similar to studies in females (Friedmann, 2002; Stoker et al., 2000; Trentacoste et al., 2001). However, Stoker et al. (2000) noted
that decreases in epididymal weights measured in higher-dose rats were no longer significant when body weight decreases were factored in as covariates. Atrazine also caused delay in puberty and reproductive tract development (lowest-observed-adverse effect 12.5 mg/kg/day, NOAEL 6.25 mg/kg/day) as well as reduced serum and intra-testicular testosterone levels, both acutely (from postnatal day [PND] 46 to 48) and chronically (from PND 22 to 48) (Friedmann, 2002). Stoker et al. (2000) postulated that the mode of action for delaying puberty in the male rat was through alteration of the secretion of steroids and subsequent effects on the development of the reproductive tract, which appear related to atrazine’s effects on the CNS.

Serum triiodothyronine (T3), thyroxine (T4), and thyroid-stimulating hormone (TSH) were unaltered by atrazine in the female pubertal study (Laws et al., 2000), which was consistent with no noted histopathological/morphological changes in the thyroid. No differences were observed in TSH and T4 between the atrazine-dosed male rats and the control group; however, T3 was elevated in the high-dose (200 mg/kg/day) treated group (Stoker et al., 2000). No effect on thyroid histopathology or hormone levels was detected in other male or female pubertal-type assays performed to date (Laws et al., 2003; Stoker et al., 2002).

**Fish short-term reproduction.** Battelle (2005) evaluated the effects of atrazine on fathead minnows (Pimephales promelas) using an earlier version of the Fish short-term reproduction assay during EPA’s validation of this test. Results indicated that atrazine exposure at levels as high as 223 µg/L had no significant effect on important assay endpoints. Some trends in the data suggested that atrazine exposure did cause more subtle effects as follows: In females, atrazine exposure at both the low and high treatments caused a slight decrease in estradiol levels, whereas in males, atrazine treatment at both the low and high exposure levels lowered circulating testosterone and 11-keto-testosterone levels by approximately 30–50%. In addition, some effects on testicular histology were noted, but the biological relevance of these histological changes was unclear because no other histological abnormalities were observed.

In another short-term (21 days) reproduction assay using fathead minnows, Bringolf et al. (2004) observed decreasing trends in relative testis weight, testis maturity, and percentage embryo fertilization, but the differences in these and other endpoints were not statistically significant in the atrazine-exposed fish. Comparison of nearly all endpoints measured in positive control fish treated with estradiol (0.5 µg/L) to those of atrazine-exposed fish and control fish showed significant differences, suggesting that atrazine did not have strong estrogenic effects in adult fathead minnows and did not cause overt reproductive toxicity at environmentally relevant concentrations.

In a review of literature studies, Eldridge et al. (2008) state that atrazine exposure has not been shown to affect vitellogenin in goldfish (Carassius auratus) or carp (Cyprinus carpio) or induce vitellogenin messenger RNA in zebrafish (Danio rerio). A recent meta-analysis was performed to extract trends from existing literature and found some consistencies for freshwater vertebrates (Rohr and McCoy, 2010). Atrazine consistently reduced growth rates, had variable effects on timing of metamorphosis that were often non-monotonic, reduced immunity, and induced diverse morphologic gonadal abnormalities associated with modified levels of sex hormones. However, in no study did atrazine affect levels of vitellogenin, implying that atrazine does not function via the ER in fish.

**Level 4 WoE** in fish tests confirms a lack of estrogenic activity. In mammals, atrazine appears to affect steroidogenesis indirectly via the CNS and by inducing aromatase activity and possibly by suppressing other genes involved in steroidogenesis (Higley et al., 2010). Therefore, the primary mechanism of action affecting E, A, and T hormone systems has been elucidated, and lowest and no-effect levels established for most of the endpoints observed in these studies. A lack of estrogenic activity in birds and a lack of E, A, or T activity in amphibians are indicated as well. Further testing is recommended only if required for regulatory purposes and should focus on effects in mammals.

**Level 5: Reproductive/Developmental Toxicity In Vivo Testing**

In recent studies looking at early developmental effects, exposure during gestation and early postpartum (via the mother’s milk) to atrazine at 100 mg/kg/day resulted in delay of preputial separation and affected the prostate in adult Long-Evans rats (Rayner et al., 2007). Exposure to atrazine during gestation from PND 14 to parturition resulted in decreased pup survival (10 mg/kg/day and above), decreased anogenital distance (75 mg/kg/day and above), and delayed preputial separation (at 50 mg/kg/day and above) (Rosenberg et al., 2008). Atrazine exposure did not affect testosterone levels in newborn pups’ testes; however, serum testosterone levels were significantly reduced at PND 60 (50 mg/kg/day and higher). According to the authors, “These results, taken together, are suggestive of antiandrogenic effects of gestational atrazine exposure on male offspring, though these effects occur at doses that are unlikely to be experienced under any but experimental conditions.”

In a developmental toxicity study with SD and Long-Evans rats, oral administration of atrazine in doses up to 200 mg/kg/day on gestation days 6–10 resulted in full-litter absorptions (none seen after the LH-dependent period of pregnancy), and delayed parturition. This suggests that effects were maternally mediated and consistent with loss of LH support of the corpora lutea (Narotsky et al., 2001).

WoE following consideration of Level 5 information leads to the conclusion that atrazine delays puberty and sexual development in both male and female rodents and has long-term effects in adult male testes. Although it does not function
through either the estrogen or androgen receptor, atrazine does affect estrogen and androgen pathways indirectly via suppression of GnRH and by modulating steps within the steroidogenesis pathway. Atrazine does not affect thyroid hormone–dependent processes in rodents or in amphibians (Xenopus laevis). Atrazine does not appreciably affect development or sexual differentiation in amphibians or fish.2

CONCLUSIONS

By using a multilevel testing framework and applying an iterative WoE analysis at each level, a number of tests required under the EDSP can be avoided, saving money and animal lives. For a chemical such as atrazine, all of the Tier 1 assays with the possible exception of the male and female pubertal assays could have been avoided, which would save more than 400 animals and over $260,000 in direct assay costs (Table 1, based on median EPA estimates [OECD, 2010]). The tests listed in Table 3 and Figure 1 provide a snapshot of currently available methods; however, the decision framework is designed to account for evolving methodology. Methods can be added to or removed from the various levels, depending on the nature of the endpoints assessed. For example, it is expected that additional assays addressing thyroid and other endocrine pathways will be added to Level 2 in the near future. With advances in the application of in vitro metabolizing systems, it is conceivable that Level 3 tests could become obsolete.

In addition to streamlining and expediting the assessment of individual chemicals, the approach described here can greatly increase the number of chemicals characterized each year with regard to their ability to affect the endocrine system, while, at the same time, potentially saving tens of thousands of animals from being killed.

SUPPLEMENTARY DATA

Supplementary data are available online at http://toxsci.oxfordjournals.org/.

ACKNOWLEDGMENT

The authors would like to thank Jabeen Akhtar for critical editorial assistance.

REFERENCES


