Pure & Appl. Chem., Vol. 70, No. 9, pp. 1805–1827, 1998.
Printed in Great Britain.

Natural and anthropogenic environmental oestrogens: the scientific basis for risk assessment*

Review of suggested testing methods for endocrine-disrupting chemicals

S. Jobling

Department of Biology and Biochemistry, Brunel University, Uxbridge, Middlesex, UB8 3PH, UK

Abstract: Recent concern about the possible impact of exposure to endocrine-disrupting chemicals (EDCs) on humans and wildlife has led to a need for the development of regulatory test methods to facilitate identification of endocrine- disrupting substances and their effects, both at the stage of product development, and when they are present in the environment. Whilst structural activity relationships (SARs) and *in vitro* tests have potential utility for the rapid identification of suspect chemicals, they do not accurately mimic effects in whole animals and are therefore complementary to, not substitutes for, *in vivo* tests on whole animals. A tiered structure for testing has been suggested by scientists at various workshops and is reiterated here. Prioritisation screens involving *in vitro* and *in vivo* short-term tests should be followed by partial or whole life-cycle studies on whole animals with a variety of reproductive and developmental end-points. Whilst existing *in vivo* mammalian test methods are broadly suitable as screens for assessing potential endocrine-disrupting effects in mammalian wildlife, it is uncertain if these assays would be of use as screens for other classes of vertebrate wildlife, due to differences in endocrine function. Existing full and partial life-cycle tests with some avian and fish species could also identify endocrine disrupters; however, these long term tests are not suitable for routine screening without modification. Several non-regulatory tests are suitable for development and could be applied for regulatory purposes after modification and standardisation. Despite the absence of properly validated internationally agreed test methods, several countries have already taken action to restrict or prevent the use/discharge of certain endocrine-disrupting chemicals.

INTRODUCTION

The potential for man-made chemicals to interfere with reproduction and development in wildlife was first recognised in the early 1960s, when the feminizing effects of the pesticide DDT were first reported [1]. Indeed, as early as 1938, it was found that certain alkylphenolic chemicals could bind to the oestrogen receptor [2]. In consideration of these facts, it may seem surprising that more than thirty years later, scientists are holding numerous workshops [3–5], both nationally and internationally, in an attempt to develop strategies for testing new and existing chemicals for endocrine-mediated effects. Indeed, whilst the concept of endocrine-mediated toxicity is certainly not new to reproductive toxicology, the historical development of existing tests and procedures has been a reactionary process and hence, relatively few of the 100,000 chemicals in existence, or their metabolites and/or degradation products, have been adequately tested. Regulatory reproductive toxicity tests for pesticides, for example, were developed largely in response to the effects of DDT seen in wild species of birds [6,7], whilst the effects of drugs such as diethylstilbesterol (DES) in humans gave rise to the existing developmental toxicity tests

Pure & Appl. Chem.*, 1998, **70(9)—an issue of special reports devoted to Environmental Oestrogens.

for therapeutic oestrogens and progestrogens. In addition, although comprehensive testing protocols for reproductive toxicity exist, there is now vigorous debate over whether these test protocols can adequately detect endocrine-disrupting effects [8]. Similarly, the significance of endocrine-related effects may not have been noted or reported within the strictures of a particular test protocol, and therefore have not featured within the overall assessment of the toxicity of a particular substance. Although no-observed effect concentrations are reported in test regimes, the effects that are assessed may not have included endocrine-related end-points. In addition, these toxicity tests use considerably higher doses, and different exposure routes than would be expected in an environmental situation.

Chemicals which can interfere with reproduction and development have been variously termed reproductive or developmental toxicants, endocrine disrupters (EDCs), or endocrine modulators [4,5,8]. To avoid confusion therefore, the term 'endocrine disrupter' will be used throughout this review to refer to a man-made or naturally occurring substance that can mimic, or interfere with, the biosynthesis, binding and/or action of natural hormones, and thereby disrupt physiological processes which are under hormonal control. Endocrine-disrupting chemicals may, therefore, also be classified as reproductive toxicants, although it must be appreciated that reproductive toxicity may occur as a result of changes caused by toxicity in other systems, in which case the effects seen may reflect other modes of action not involving the endocrine system.

In wild populations of animals, almost all of the historical instances of endocrine effects that may have been caused by exposure to EDCs have been manifest as disruption of reproduction and/or sexual development and differentiation [9]. Consequently, there has been an understandable emphasis in research effort on chemicals that act via interaction with nuclear hormone receptors, particularly via the sex-steroid hormone receptors. Whilst other mechanisms of action are obviously possible, this review will also focus largely on strategies and procedures for testing chemicals with sex-steroid hormone receptor binding properties, and to a lesser extent, on the general principles of testing strategies for chemicals which might interact with other members of the nuclear hormone receptor family.

SEX-STEROID HORMONE MIMICS AS MODEL ENDOCRINE DISRUPTERS

Because of the important role played by sex-steroid hormones during development, xenobiotics with hormonal action can potentially disrupt a variety of developmental processes, particularly those related to the sexual differentiation and development of the reproductive tract and central nervous system [10,11]. Indeed, although exposure to relatively high doses of known hormone mimics may be necessary to induce marked disruption in adults, it is possible that considerably lower levels may be harmful if exposure occurs over long periods of time and/or at a critical time in early development. These effects, although more subtle than those caused by exposure to a carcinogenic or neurologic environmental toxin, could affect reproduction and development, and ultimately the survival of exposed species. Evidence from humans where mothers who took DES during pregnancy, for example, suggested that permanent adverse effects in their offspring could be induced by very small doses of the drug at critical times during the development of the foetus [12,13]. Similarly, in rats, methoxychlor exposure throughout gestation also results in reduced testes size and lower sperm counts in the offspring [14]. These critical windows of development in mammals represent a period of heightened vulnerability to endocrine-disrupting chemicals. Indeed, the extreme sensitivity of the foetus to the hormonal environment in the uterus is illustrated by the fact that natural differences in hormone levels surrounding mouse foetuses of only 10^{-12} M influence the timing of sexual development and the behaviour of the animal in adult life [15,16].

The early life stages of other animals also appear to be especially sensitive to sex-steroid hormones, and these compounds have been shown to have profound effects on sexual differentiation. As an example, in fish, exposure to exogenous oestrogen during a narrow window spanning 10 days either side of hatching can cause feminization of the subsequent fry [17]. Similarly, exposure of juvenile fish to the xeno-oestrogens 4-pentylphenol or 4-nonylphenol induces hermaphroditism and may even result in complete feminization of males [18,19]. In oviparous animals, lipophilic endocrine disrupting chemicals may concentrate in the developing oocytes because the yolky eggs sequester large quantities of lipid. Furthermore, because ovarian development in these animals involves the mobilisation of lipid reserves from the maternal stores, this may 'free-up' lipophilic compounds that have been bioaccumulated in the

fatty tissue in the adult. In turn, these compounds may be transported with the lipid into the developing oocyte; this has been shown to occur with o,p′-DDT and PCBs in the Atlantic croaker, *Micropogonias undulatus* [20]. Eggs therefore, may contain high concentrations of endocrine disrupters. Subsequently, the utilisation of these lipid reserves by the developing embryo would release xenobiotics to the embryo at a very sensitive time when the sex is labile [21]. The view that hormones (or their mimics) could not influence embryonic development until the endogenous hormone itself was present no longer holds true, as embryonic cells may exhibit receptors before the hormones themselves are synthesised [22,23]; as an example the ER is found in embryos as early as the 8-cell stage [24].

Sex steroid hormones also play a major role in reproduction and sexual development and consequently, EDCs could influence both of these processes. For example, an oestrogen agonist could interfere with the feedback inhibition of steroidal synthesis, leading to increased steroidogenesis in the ovaries or testes. Moreover, decreases in plasma androgens have been shown to accompany exposure to oestradiol in several fish species, presumably due to feedback inhibition of androgen synthesis [25]. Steroids are also important in the control of secondary sexual characteristics; for example, both breeding tubercles in some fish species, and sex-specific colour changes serve as good indicators of breeding condition. Oestrogens and androgens can also act directly on the secretory activity of the pituitary gland and exert both positive and negative effects on gonadotrophin release [26,27]. EDCs that mimic or antagonise the action of steroid hormones could therefore disrupt reproductive function by altering the secretion of gonadotrophins [28].

OTHER NUCLEAR HORMONE RECEPTOR MIMICS

The sex-steroid hormone receptors belong to a nuclear hormone receptor super-family which also includes the retinoic acid, thyroid hormone, prostaglandin and vitamin D receptors. The ligands for these receptors are also important in reproduction and/or development. Thyroid hormones, for example, play an important role in the differentiation of the Sertoli cells in the developing testes in mammals [29], whilst in amphibians, they are important for metamorphosis. Similarly, the retinoic acids have important functions during embryogenesis and early development [30] and therefore, xenobiotics that are able to bind to the retinoic acid or thyroid receptors might affect growth and/or reproduction or induce severe morphological defects. Indeed, it has been known since the early 1970s that the retinoic acids 9-*cis* retinoic acid, all-*trans*-retinoic acid, 13-cis retinoic acid and 9-cis-13-cis-retinoic acid are vertebrate teratogens, producing heart, cranofacial, and limb malformations. Similarly, Harmon *et al.* [31] showed that both the insecticide s-methoprene and its hydrolytic degradant, all-trans-S-methoprenic acid, bind to the retinoic X receptor, RXR, through mimicking the structure of natural retinoic acids. Much more recently, using amphibian embryos, La Clair *et al.* [32]were able to convincingly show that, whilst S-methoprene itself exerts no adverse effects, its natural photo-and hydrolytic degradants are embryotoxic and cause malformations in brain development, misshapen eyes, poorly coiled guts and oedema. Retinoic acids are also important in embryogenesis in all other vertebrate groups, and therefore xenobiotics that mimic or antagonise the action of these hormones could potentially affect other animals, not just amphibians.

During the last 5–7 years, the number of chemicals known to have endocrine-disrupting potential has increased exponentially. In addition to pesticides and pharmaceuticals, other groups of chemicals, such as the alkylphenolic chemicals(breakdown products of industrial surfactants) [18,33,34], have also been found to cause reproductive and developmental effects in wildlife. These factors, together with the wider appreciation that many of these chemicals are/could be present in the environment in quantities that are sufficient to cause endocrine effects in wildlife populations [35], have led to a consequent increase in the level of concern about the possible impact of exposure to such chemicals on human and environmental health. In particular, the widespread occurrence of intersex fish in British rivers [36] and the apparent decline in human male sperm counts [37–39] in some countries have been hypothesised to be linked to exposure to endocrine-disrupting chemicals (particularly oestrogens) that are present in the environment. Much more recently, the widespread occurrence of deformities in wild populations of amphibians in the USA has been associated with developmental exposure to the degradation products of s-methoprene, an insecticide used to control fleas and mosquitoes [32].

EDC SCREENING METHODS…THE CURRENT STATUS

There are presently no rigorous guidelines for the assessment of environmental health in wildlife populations which may be affected by exposure to EDCs. Consequently, the assessment of a novel chemicals endocrine-disruptive activity is heavily reliant on routine regulatory toxicology testing. Although individual countries have specific testing requirements, the guidelines produced by the Organisation for Economic Co-operation and Development (OECD) and the US Environmental Protection Agency (EPA) are generally representative; relevant designs including mammalian general developmental and reproductive toxicity studies and a number of simple studies on non-mammalian species. Perhaps the most comprehensive tests, with respect to endocrine disruption, are those which assess *in vivo* mammalian toxicity, including some which specifically address reproductive aspects. These types of test are carried out using laboratory rodents and enjoy wide application in many toxicology laboratories. Typically, they employ a wide range of end-points in both the adult and, in some cases, also in the offspring [40]; examples of these tests are shown in Table 1.

The most detailed tests include exposure of the adults prior to mating and throughout gestation, followed by examination of sexual differentiation and gonadal development in the offspring. Although they are obviously intended to provide data for human risk assessment, they also provide data of potential use in assessing effects in other mammalian species. Indeed, if properly validated, it may be that some mammalian tests can be extrapolated into effects on other groups of vertebrates. In most cases, however, these existing tests need to be enhanced by the inclusion of more suitable end-points and/or changing the exposure period, in order to improve their ability to detect endocrine-disruptive activity(see later discussion of *in vivo* tests).

The enhancement of existing test protocols is also required in avian reproductive toxicity tests in order to improve their ability to detect endocrine disruption. Birds were one of the first classes of wildlife in which endocrine disruption, thought to be due to the pesticide DDT, was first reported [6,41,42], and hence avian tests to assess reproductive capabilities are a usual requirement for pesticide registration. Reproduction is assessed using pairs of animals and the collected eggs are examined for shell thickness, viability and hatching success [43]. As they are currently performed, these tests are unsuitable for the assessment of EDCs, primarily because the protocols do not comprehensively evaluate the F1 progeny or the second generation in which more adverse effects could occur and/or in which effects could occur at lower concentrations.

In almost all other groups of animal, adequate tests for EDCs simply do not exist. In fish, for example, both the OECD and the US EPA require *in vivo* tests for various regulatory purposes, including short-term lethality, early life-stage and partial and full life-cycle studies. Commonly tested species include carp, guppy, rainbow trout, sheepshead minnow, fathead minnow, medaka and zebrafish. Although most of the end-points used in, for example, the early life stage test [44], would detect the effects of chemicals which may cause gross morphological abnormalities, none would detect oestrogenic or androgenic effects on sexual differentiation. Furthermore, basic reproductive toxicity tests in adult animals would be more useful as screening tools if additional, more informative, end-points were included in the protocols. Measurement of, for example, plasma vitellogenin concentration in male or immature fish would facilitate the detection of oestrogenic activity. Moreover, if it were established that vitellogenin production is indicative of a more serious biological effect, then this test could be used to prioritise chemicals for further testing, or eventually replace the longer, more costly, partial life cycle tests.

In reptiles and amphibians, no statutory requirements for testing exist. Only the frog embryo toxicity test (FETAX) has been standardised and validated in several laboratories [45]. The end-points of this test include survival, growth and malformation of embryos and thus, as it is designed to examine teratogenicity of chemicals to frog embryos, this test could be used to screen chemicals for developmental effects that may be mediated by the retinoic acid receptors [32]. It is also possible that similar tests, involving exposure of tadpoles, could be modified to assess endocrine disruption by oestrogen receptor agonists/antagonists and/or thyroid hormone ligands.

Table 1.

Test	Dosing regimen	Shortfalls	Endocrine end-points
Multigeneration reproduction and/or continuous breeding studies (rats and mice)	Repeated administration throughout the reproductive cycle of males and females	With the addition of the proposed modifications, there are no shortfalls. This test will, however need further validation if it is to be used to predict the likelihood of effects in non-mammalian wildlife species.	insemination, fertility rate, fecundity, perinatal development over one or more generations. New proposed EPA test will also incorporate sperm count, motility and morphology. Oestrous cycle length, developmental landmarks such as vaginal opening, detailed histopathology of the ovaries, testes and epididymides
Prenatal developmental toxicity studies (rats, mice, rabbits)	Single or repeated administration during organogenesis of foetuses in pregnant animals. New proposed EPA test will incorporate exposure from implantation to full term.	With the proposed modification, this test is now suitable to detect effects which might be manifest during Sertoli cell differentiation. However, end-points do not include examination of fertility, gonadal development and histopathology in the offspring.	Implantation, abortion, early embryonic death, external and internal malformation and anomaly rate.
Avian reproductive toxicity test (Mallard or Japanese Quail)	Exposure via the diet for 20 weeks through to egg hatch.	Suitable for higher tier testing, beyond the prioritisation and screening level although additional end-points(e.g. body weight, gonad weight, gonad histopathology and gross pathology, sex ratio) should be examined in the offspring. Sacrifice of offspring should be at day 8–10 post hatch in order to see effects on sexual differentiation	Mortality of adults, egg production, cracked eggs and egg shell thickness, viability, hatchability, effects on young birds up to 14 days post hatch
Fish Early Life Stage Test	Exposure from fertilisation through to juvenile stages	Potentially useful with modification to include histopathology, sex ratio and sex differentiation	Hatchability and viability of eggs, growth and development of larvae including abnormal behaviour and appearance
Fish Full Life Cycle Test	Exposure from fertilisation through to 8 weeks post hatch in the second generation	No histopathology or biochemistry end-points included. Very costly.	Viability of eggs, development, growth, reproductive success of second generation

In summary, the current scientific consensus is that there is a need to develop and/or modify hazard identification approaches that are able to identify substances that may cause endocrine disruption in wildlife. These tests and procedures should be optimised and validated prior to regulatory acceptance, and should facilitate identification of endocrine-disrupting substances and their effects, both at the stage of product development, and when they are present in the environment. Any new or revised test must satisfy a number of minimum criteria if it is to be used for regulatory purposes. These include the demonstration of a link between the new test and effects in a target species, the generation of useful data for hazard identification and risk assessment, adequate validation and standardisation both between and within laboratories using representative types of chemicals, and cost effectiveness. Whilst existing mammalian reproductive and developmental toxicity tests are, in many cases, broadly suitable for the assessment of reproductive toxicity of existing chemicals, more comprehensive testing needs to be

applied to other vertebrate groups also. In addition, the inclusion of specific end-points in *in vivo* assays which are indicative of the various known mechanisms of action for endocrine disrupters may prove useful, as would the development and validation of *in vitro* assays that determine, for example, receptor binding and/or activation. Indeed, although it has never been necessary to provide specific information on mechanisms of action to regulate chemicals as reproductive toxicants, a full knowledge of the mechanism of action would facilitate the future design and synthesis of new chemicals which do not disrupt the endocrine system. To illustrate this point, consider a well-known historical example of endocrine disruption, in which exposure of female dog-whelks to low concentrations of tributyl-tin (TBT) in the marine environment, caused widespread masculinizing effects [46,47] and resulted in sterility and local extinction of mollusc populations in contaminated areas. Although TBT was justifiably banned several years ago without any knowledge of its mechanism of action, this knowledge [48] could now be used to develop *in vitro* tests to screen chemicals for their ability to act via this mechanism. Indeed, in cases where the mechanism of action is known or suspected, such tests, could be much less costly and labour intensive than *in vivo* tests on whole animals. The development of appropriate *in vitro* tests would, however, involve their validation *in vivo* prior to their routine application (see later discussion on the relative merits of *in vivo* versus *in vitro* tests). A further important consideration concerns the fate and behaviour of chemicals in the environment: There are now several documented examples of seemingly 'harmless' chemicals which are rapidly converted into more detrimental materials, via metabolism or degradation [32,49–51]. Similarly, chemicals that lose their hormonal activity through degradation also exist. These examples suggest that the criteria for analysing the environmental impact of a chemical should be extended to consider carefully the products of its metabolism and natural environmental degradation in order to minimise future environmental impact.

In nature , wildlife are usually exposed to mixtures of chemicals and their degradation products (e.g. sewage treatment work effluents), and hence in most cases, it is not clear which chemical is causing adverse effects. Although some attempts have been made to investigate the interactions between endocrine-disrupting chemicals when used in combination [52,53], few of these studies have tested the individual substances at equimolar concentrations to that of the combined dose and therefore, the results are difficult to interpret. It is a well known fact that combinations of substances can lead to enhanced activity over that of individual substances. Indeed, additivity in other toxic reactions (e.g. narcosis) is a recognised response where substances have similar modes of toxic action. There is also some evidence to suggest that the metabolic pathways involved in the breakdown of some chemicals are disrupted by exposure to other xenobiotics simultaneously[54]. The interactions between various realistic combinations of endocrine-disrupting chemicals (for example, nonylphenol, nonylphenol ethoxylates and nonylphenol carboxylates), are poorly understood. Consequently, there has been a growing awareness that, despite the many difficulties, studies of interactions of chemicals with each other and the effects of realistic mixtures of chemicals on organisms are required.

SUGGESTED STRATEGY

A tiered strategy for testing chemicals has been suggested by scientists at various recent workshops [3–5,8]. This should begin with short-term *in vivo* and/or *in vitro* tests with endocrine end-points, in order to prioritise chemicals for further *in vivo* testing in chronic short-term assays, with reproductive and developmental end-points, such as output of the gametes, or sexual differentiation of the young. The third tier would be long-term whole life cycle studies covering one or two generations with similar end-points to the chronic short-term assays. It is important to appreciate that, due to the complexities of the endocrine system, any screening design should not consist of a single end-point, but rather a suite of end-points that are reflective of the nature of the concern (e.g. oestrogenic, anti-androgenic etc.). Furthermore, in order to facilitate the development of standardised screening protocols and inter-laboratory validation, these end-points should be relatively easy to measure and minimise the occurrence of false negatives or positives. In this regard, *in vivo* tests on whole animals appear inherently more suitable than *in vitro* tests because they minimise uncertainties related to differences in xenobiotic metabolism, bioavailability, and toxicokinetics between *in vitro* and *in vivo* systems. Notwithstanding this, these same features may render the results of *in vivo* assays less useful for extrapolation between

different species/vertebrate groups. In addition, the widespread application of *in vivo* testing could prove to be too resource intensive to evaluate the thousands of chemicals which are of potential concern. Indeed, the need for the development of rapid prioritisation screens, such as those which can be conducted *in vitro*, becomes obvious when one considers that over 100 000 chemicals are listed on the European inventory of Existing Chemical substances. In addition there are 1500 to 2500 new industrial chemicals (excluding pharmaceuticals and pesticides) submitted yearly in the USA for evaluation under the Toxic Substances Control Act; it would probably be impossible to screen all of these chemicals *in vivo*. It is clear, however, that *in vitro* tests and predictive models based on chemical structural similarities are intrinsically unable to perform the screening function on their own presently, due to the sheer complexity of the endocrine system. However, information on a chemical and its mode of action would clearly be of considerable use in the decision about which combinations of tests should be used first. It could even be possible to proceed directly to more longer term tests, if sufficient information on the mode of action is already known. In consideration of all of these facts, it is clear that *in vitro* tests, if properly validated- *in vivo* , could have great utility at the initial stage of prioritisation screening.

STRUCTURE ACTIVITY RELATIONSHIPS (SARs) AS PREDICTIVE TOOLS TO IDENTIFY EDCS

Structure Activity relationships (SARs) are based on the principle that the properties and behaviour of chemicals are derived directly from their molecular and structural characteristics. Specifically SARs describe the chemical and/or biological properties of a series of chemicals relative to their molecular structure and/or other physicochemical properties [55]. Once such a relationship is established, it can be used to predict the activity of untested chemicals. Consequently, SARs could serve as highly effective tools for screening and prioritising possible EDCs for further, more detailed investigation.

It was first thought that it was not possible to predict the hormonal activity of chemicals based on a knowledge of their structures, as not all oestrogenic chemicals, for example, appeared to have structures which superficially resembled that of oestradiol (e.g. kepone). This difficulty, however, may, at least in part, be attributed to the mix of chemicals in any one formulation. To illustrate this point, if chemical x is, say, 98% pure and has 1/10 000th the potency of oestradiol-17β and the impurity has 1/200th the potency of oestradiol-17β, then the impurity would account for the entire activity of x. Furthermore, most industrial chemicals are complex mixtures of isomers with differing potencies and modes of action [56]. It is now clear that some (probably most) xeno-oestrogens produce their oestrogenic effects by binding to the oestrogen receptor in an identical manner to oestradiol itself [57] and the structures of most chemicals that are known to mimic oestradiol do appear to be based around one or more aromatic rings. Structure-activity studies have been carried out on steroidal and non-steroidal ligands to several members of the steroid receptor family, including the oestrogen receptor - alpha and beta [58,59], and androgen receptor [60], amongst others. The most extensive study to date in terms of a wide variety of steroid and non-steroidal ligands was reported by Waller *et al.* [58] for ER binding affinity. In this study, steric and electrostatic properties of 55 compounds, representing oestrogens, androgens, progestrogens, phyto-oestrogens, diethystilbesterol derivatives, organochlorine insecticides, PCBs and their hydroxylated metabolites, phthalates, and phenols were related to binding affinity in a statistically robust and internally consistent manner. It is becoming increasingly evident, therefore, that some central structural features of xeno-oestrogens have to exist for them to be oestrogens, if they bind to, and exert their effects, through the oestrogen receptor. It is likely that this is also true of chemicals that mimic other hormones by acting through the appropriate receptors. Ultimately, predictive models describing the potential hormonal and anti-hormonal activities of all man-made chemicals will be required. The development of such models will require the incorporation of receptor-binding affinity data and information on the type of hormonal effect which the chemical displays, both *in vitro* and *in vivo*.

There is some optimism that SARs will be available in a relatively short time frame(3 years), at least for a limited number of receptors and restricted classes of chemicals. It should be remembered however, that initial expectations of the use of SARs in the identification of EDCs may prove to be overly optimistic. Their limitations, with respect to their predictive power, are due to the shortfalls in the *in vitro* and *in vivo* systems used to generate the data. In general, SARs are based on *in vitro* data, and therefore

extrapolation to an *in vivo* situation can easily lead to false negatives for compounds which are bioactivated *in vivo*, or to overestimates of potency for compounds readily degraded *in vivo*. The two metabolites of vinclozilin, but not the parent compound, for example, bind to the androgen receptor resulting in the anti-androgenic activity of vinclozilin seen *in vivo* [51]. Similarly, the photolytic degradation products of the insecticide s-methoprene, but not the parent compound, bind to the retinoic acid receptor [31] and result in severe morphological deformities in amphibian embryos [32]. In addition, it is clear that mere binding to the receptor will not be indicative of whether the compound will act as an antagonist or an agonist. There are, however, methods to address this issue and a SAR prediction of receptor binding could be used as a trigger to require that such an assay be performed [56]. Other chemicals with endocrine-disrupting properties that would be difficult to detect using nuclear hormone receptor SARs include those chemicals that do not bind to receptors. The occurrence of imposex (masculinisation of female animals) in some marine gastropod species by TBT (which does not bind to the ER or AR), for example, is due to inhibition of aromatase. This inhibition results in a limited conversion of testosterone to oestradiol [48] and results in the formation of a penis in females. Other factors which may lead to uncertainties in SARs include inter-species differences in receptor affinity and multiple receptors in a particular hormone receptor system. Recent reports, for example [61], have identified multiple isoforms of the oestrogen receptor in the rat; the ER alpha and ER beta are very different in terms of their tissue distribution and affinity for ligands. There are also a number of uncertainties that are directly related to the different SAR modelling approaches, rather than the development of *in vitro* and *in vivo* assays in general. In particular, it is absolutely crucial that the training-set of chemicals used to calibrate the model must be representative of the chemicals for which the predictions are to be made. This can be achieved by systematic selection of the training set of chemicals among a larger group of chemicals, where the major structural features which characterise these compounds are varied systematically and simultaneously.

In summary, SARs have greatest potential as a prioritisation screen, to assess large numbers of chemical structures simultaneously and indicate structural features which may indicate their ability to interact with hormone receptors. The identification and subsequent testing of natural degradation products and metabolites would greatly enhance their applicability and widespread use. It is unlikely, however, that they can perform the screening function on their own due to the sheer complexity of the endocrine system of intact organisms.

THE POTENTIAL USE OF *IN VITRO* BIOASSAYS

The incorporation of *in vitro* tests into the assessment of the impact of endocrine modulating chemicals on wildlife species offers a number of advantages and is consistent with the guidelines set-forth in the EEC Directive 86/609/EEC which states that, 'an experiment shall not be performed if another scientifically satisfactory method of obtaining the result sought, not entailing the use of an animal, is reasonably and practically available', and hence emphasises the fact *in vivo* tests should be replaced with *in vitro* tests wherever possible. Several *in vitro* assays for measuring oestrogen/anti-oestrogen, androgen/anti-androgen and dioxin-like activity of individual compounds or complex mixtures have been developed. These assays use a variety of end-points, including enzyme and gene induction, ligand binding, increased protein expression and cell proliferation and differentiation [62]. It is noteworthy however, that none of these tests have ever been applied in a regulatory manner. The advantages and limitations of the various *in vitro* screens have been reviewed in detail elsewhere [62]. In the following section, several of these assays will be briefly discussed . Although these examples describe assays used to implicate substances as xeno-oestrogens, the same principles can be ascribed to any chemical which exerts its action through a nuclear hormone receptor.

COMPETITIVE RECEPTOR BINDING ASSAYS

Competitive binding assays for the oestrogen receptor have been extensively used to investigate the interactions of various ligands with the ER [52,57]. Typically, the ability of increasing concentrations of a test compound to displace radiolabelled ligand from the receptor is assessed and hence, although these

assays cannot distinguish between receptor agonists and antagonists, they are useful in providing an initial assessment of whether a suspect chemical or its metabolites/degradates are likely to bind to the ER. It should be remembered, however, that mere binding to the ER is not sufficient to determine oestrogenicity of a substance, since potency is dependent on binding affinity and the ability to initiate or inhibit transcriptional activity and hence, cause an adverse response. Moreover, high concentrations of a competitor ligand may result in non-competitive displacement. At these high concentrations (in the micromolar range), many of the xeno-oestrogens tested may also approach the limits of solubility, thus making the interpretation of any results sometimes difficult.

CELL PROLIFERATION ASSAYS

Traditionally, assays involving the proliferation of ER-positive, oestrogen-responsive MCF-7 or t47-D human breast cancer cells have been widely used for the assessment of oestrogenic substances. It is a well-known fact that oestrogens induce cell proliferation of human oestrogen-sensitive cells by negating the inhibitory effect of a cell proliferation repressing factor. The 'E-Screen' [63]evaluates the number of cells present following a 6-day incubation period in medium supplemented with steroid-stripped serum in the presence or absence of a suspect xeno-oestrogen. Cell proliferation is measured using a Coulter Counter or haemocytometer or by determining the level of tritiated thymidine incorporation. The reported detection limit of 10 pgE2/mL (30 pM) makes the MCF-7 cell E-Screen one of the most sensitive assays for assessing oestrogenicity of EDCs. Many factors have, however, been shown to influence the relative potency of oestrogenic substances, including differences between cell line clones, culture conditions, receptor level differences, differences in serum and cell density (see Zacherewski [62] for a comparison between studies). In addition, the level of oestradiol necessary for optimal growth stimulation has been reported to range from 10 nM to 10 pM whilst the level of induction ranges from 0.4 to 8.0-fold! [64]. It must also be noted that, to date, several mitogens have also been found to enhance the proliferation of human breast cancer cells [65]and therefore caution must be exercised when using this assay alone to screen chemicals for oestrogenic activity. Further disadvantages that would affect regulatory acceptance of this screen include the lack of equivalent cell lines with proliferative end-points in wildlife species and therefore the questionable predictiveness of *in vivo* effects in wildlife.

PROTEIN PRODUCTION ASSAYS

The induction of several proteins by various types of cultured cell have also been used to assess oestrogenic potency of suspect chemicals. Examples of these include induction of the progesterone receptor and increased expression of pS2 in breast cancer cells, prolactin production by immature rat pituitary cells and induction of vitellogenin production by fish liver cells [66–69]. The end-points employed in such assays are highly specific and sensitive and also allow distinctions to be made between agonist and antagonist. In addition, the use of a wide variety of cell line origins enables the investigation of agonist and antagonist activities in different potential target tissues, which may be an important factor when extrapolating between tissues and species. Tamoxifen, for example, exhibits anti-oestrogenic activity in breast tissue and cells but is weakly agonistic in the uterus and endometrial cells [70]. Cross-species differences are seen with the human anti-progestin RU486 which is inactive in chickens and hamsters because it does not bind to the receptor due to a single point mutation within the ligand binding domain of the chicken and the hamster progesterone receptor [71]. The major disadvantage of these screens is that the measurement of the oestrogen-induced protein often involves laborious and sophisticated techniques. In addition, the fish liver cell assay requires the fresh isolation of cells for primary culture each time it is carried out and thus, would not be suitable for widespread regulatory use.

GENE EXPRESSION ASSAYS

The construction of recombinant reporter genes consisting of promoters from oestrogen-responsive genes linked to reporter genes encoding firefly luciferase, chloramphenicol acetyl-transferase (CAT), β-galactosidase (LacZ), or alkaline phosphatase have provided probably the most sensitive, selective and

easily performed *in vitro* assays currently available. Typically, these genes are transfected into cultured animal cells containing the oestrogen receptor which are then challenged by the suspect xeno-oestrogen. Interaction and subsequent activation of the receptor causes the production of the reporter gene which can easily be measured. Various types of reporter construct can be used, the simplest of which employ endogenous promoters such as the vitellogenin complement(C3), or cathepsin D promoters. These promoters consist of several hundred base pairs of the 5' flanking regulatory region, and are, therefore, still susceptible to mechanisms of induction that are not mediated by the ER [72]. This pitfall has, however been overcome by the use of reporter genes regulated by the 13 base pair vitellogenin A2 oestrogen response element [52,57]which ensures that induction of the reporter only occurs via the ERE. Even this approach is not without complications, since other receptors for retinoids and thyroxine can interact with the ERE and thus modulate reporter gene activity [73,74]. It should be noted however, that the interaction of other receptors with the ERE is also a feature of natural oestrogen-responsive genes in the intact animal and thus, should not necessarily be regarded as a limitation of this assay. There is however a lack of availability of species-specific cell lines to enable cross-validation of this test between different species or vertebrate groups presently.

Chimeric receptor/reporter gene constructs have also been proven to have utility in screening chemicals for oestrogenic activity. For example, the E2 Bioassay [75]consists of a chimeric receptor (with ligand binding domain of the ER and the DNA binding domain of the yeast transcription factor Gal4) and a Gal4-regulated reporter gene consisting of the firefly luciferase cDNA regulated by a basal promoter and five tandem Gal4 response elements. Both of these constructs are transfected into recipient cells which are then challenged with chemical(s) under test. In this assay, oestrogenic chemicals will bind to the ER ligand binding domain of the chimeric receptor and transform the construct into an activated high affinity DNA binding receptor complex which binds to the Gal4 response element on the luciferase reporter gene. Binding of this activated complex will then initiate expression of the firefly luciferase cDNA which, in turn results in the induction of luciferase activity. Thus, luciferase activity is a direct measure of oestrogenic activity which can be detected in the pM range. Moreover, since no mammalian proteins are known to bind to this response element, increases in reporter gene activity can only occur via the activation of the chimeric receptor. Using these techniques, it should be possible to construct E2 bioassays in a variety of animals, such as fish and birds, using the ligand binding domain of their respective ER receptors. This assumes, of course, that the gene sequences of these receptors in the various species are known; ER receptor gene sequences are only known for a limited number of species in each vertebrate class. Other limitations of chimeric assays include a variation in responsiveness that may be due to differences in the quality of DNA used in the transfection, the types of cells used and/or differences in the level of expression of the oestrogen receptor [76].

YEAST-BASED ASSAYS

Gene transcription assays based on the yeast strain *Saccharomyces cerevisiae* [77] have several advantages over other known systems, including the lack of known endogenous oestrogen receptors and medium that is free of steroids that may otherwise influence the results. Indeed, due to the extreme sensitivity of these assays, oestrogenic activity can be detected down to 0.07 pM [78]. Most of the published examples of yeast based assays are based on yeast transformed with the human ER cDNA and an ERE-regulated LacZ reporter that encodes for a β-galactosidase enzyme [79]. In some instances, the galactosidase enzyme is secreted into the medium, where it reacts with a substrate and causes a colour change from yellow to red. These assays have great potential as initial screening tools for receptor-mediated effects and have been used to investigate the structural activity relationships of a wide range of alkylphenolic chemicals and organochlorines [80]. Some strains, however, exhibit agonistic activity in the presence of known anti-oestrogens such as tamoxifen and ICI 164,384, rather than antagonistic activity and are therefore of more limited use [81]. As with all *in vitro* assays, the regulatory acceptance of yeast-based screens awaits more thorough validation and optimization, as well as confirmation using *in vivo* tests.

IN VITRO VERSUS *IN VIVO*

The advantages of *in vitro* tests, in general, include cost effectiveness, speed of the test, reproducibility (consistency), capability of handling large numbers of samples and indication of the mechanism of action. Moreover, synergistic, antagonistic and additive interactions may be studied (albeit at a single site) within complex mixtures. Furthermore, in most documented cases, the results are consistent between *in vitro* and *in vivo* assays. Thus, metabolites of PCBs (hydroxyl-PCBs) are oestrogenic in receptor-binding assays, cell systems and whole animals (reptiles and birds), and alkylphenolic compounds are oestrogenic in receptor-binding assays [57], cell-systems [82]and the mouse uterine weight assay [83]. However, some chemicals can be inactive in *in vitro* assays, but active *in vivo* (such as mestranol, used in the contraceptive pill) , due to metabolism, whereas other chemicals can be active *in vitro*, but not *in vivo*, presumably due to biodegradation/metabolism. For example, several phthalate chemicals although shown to be active *in vitro* [52,84], in receptor binding assays, are inactive *in vivo* in the uterine weight bioassay [85]. These difficulties are, however, not insurmountable and could be addressed if metabolites and/or degradation products are also tested. Yet another limitation of *in vitro* bioassays is that they do not account for the effects of bioconcentration and bioaccumulation which may be important factors *in vivo* wildlife effects. In addition, they cannot accurately model the interactions involving the induction of binding proteins such as sex-hormone binding globulins that may modulate the uptake and metabolism of sex-steroids or identify critical windows within the life cycle of an organism which may be sensitive to endocrine disruption. There have also been some problems with a few of these *in vitro* techniques with respect to consistency of results between the different laboratories in which they have been employed. As an example, using MCF-7 cells, some laboratories show that some pesticides, including dieldrin and endosulfan, are oestrogenic [53], but in others they are not [86]. Similarly, using the different recombinant yeast screens , into which a gene for the human receptor has been incorporated, different laboratories have found that specific mixtures of oestrogenic chemicals can have very different potencies [87].

The assays which are briefly described here specifically screen for xeno-oestrogens that act via the oestrogen receptor. However, some endocrine-disrupting effects , such as changes in the metabolism of E2 and the effects of dioxins and some PCBs [48,88] are either mediated via different receptors, or via effects on the enzymes controlling the biosynthesis of steroid hormones. In the former case, the principles behind assays that may be suitable for screening are exactly the same as those described above and, therefore, will not be repeated. In the latter case, it should be possible to culture steroidogenic tissues in the presence of suspect chemicals and measure the concentrations of sex steroids produced by immunoassay. In many wildlife species, however, established cell lines are not available presently .

In summary, although the utility of *in vitro* assays is apparent, concerns have been expressed regarding their use in the risk assessment process. There are, however, few feasible alternatives that are capable of handling the vast numbers of samples in question. Indeed, as the list of chemicals that require testing continues to grow, the need for a means of prioritising substances via *in vitro* testing becomes apparent. Thus, although it is widely recognised that *in vitro* tests are complementary to, and not substitutes for *in vivo* tests on whole animals, several *in vitro* assays are currently being evaluated by scientists and regulators in an effort to identify those tests that can most reliably contribute to the risk assessment process. This process will, of course, involve validation using appropriate *in vivo* tests.

IN VIVO TESTS (GENERAL PRINCIPLES)

In vivo tests are more useful than *in vitro* tests in several respects; they facilitate the evaluation of mixed mechanisms of action and can be used to evaluate numerous end-points (from molecular markers to gonadal histopathology). The administration of a test substance or mixture to a sentinel/surrogate species can be carried out relatively easily to mimic an exposure scenario encountered in the environment. In addition, a whole living organism might metabolise an inactive test chemical to active metabolic products or vice versa [51,89]. In view of the above, the importance of *in vivo* screening and testing methods which are able to identify endocrine disrupters before they are released into the environment (i.e. prioritisation screens) is paramount. In addition, the development of biomarkers indicative of endocrine effects would facilitate the assessment of affected populations in the field.

Route of exposure

The basic purpose of a prioritisation screen would be to establish the plausibility of a chemical acting through a particular mechanism, that is, in terms of (anti-) oestrogenic/androgenic responses, and therefore, it has been argued that the body burden of the chemical in question may be more important than the route of exposure. In this respect, techniques such as micro injection of chemicals/mixtures into eggs or intraperitoneal injection into adults could prove useful as these methods can also be modified to screen complex mixtures of potential EDCs. It should be remembered however, that these methods can create problems; for example, if the chemical in question is toxic and causes a cessation of feeding, or a change in an endocrine effect as a secondary result of other toxicity mechanisms. Prioritisation tests should be followed up by long-term whole life cycle studies, also with reproductive and developmental end-points. Subsequently, the route of exposure should mimic that seen in the real world. In fish, for example, the principle route of exposure depends on a variety of factors, including water chemistry, the physical and chemical characteristics of the compound, and the biology of the species in question. When designing a particular test system, all of these factors should be considered. In addition, wherever possible, all tests should be accompanied by analytical support to relate effects observed with actual water concentrations or body burdens of the chemical in question.

Any study, whether for prioritisation, sub-chronic or chronic testing should be based on a traditional dose-response design. Three to five doses, which must include at least some environmentally realistic concentrations, should be used for range finding, and therefore adequate research into the known environmental concentrations of the chemical in question must be carried out.

Representative species

Establishing a minimum number of surrogate species that are representative of all other species is envisaged to be one of the biggest problems which toxicologists will encounter during the development of *in vivo* tests. The reason for this is primarily because there are substantial shortfalls in the state of scientific knowledge concerning the basic endocrinology which controls reproduction and development in most, if not all, animals. In vertebrates, the organisation of the neuroendocrine regulatory system and the hypothalamic-pituitary axis are reasonably well conserved. In addition, several of the biochemical pathways involved in the synthesis and action of steroid hormones are similar among vertebrates. Although the basic hormonal systems have been conserved during evolution, there are some clear differences and therefore it is very difficult to extrapolate with any degree of confidence between species. In male birds, for example, primary differentiation of the testes is influenced by oestradiol, whereas in mammals differentiation is under the control of testosterone. Similarly, egg-shell thinning in birds [90] could not be predicted by a mammalian or fish-based test and metamorphic effects on amphibian species could not be predicted using any other vertebrate group. Species or class differences with respect to metabolic activation or deactivation of EDCs also exist. For example, fish have lower phase 1 and 2 metabolising activity than birds and mammals [91,92] and therefore a chemical of concern might be deactivated by oxidative metabolism in mammals whilst remaining active in fish. In consideration of these facts, a screening paradigm for wildlife should ideally include *in vivo* assays with several model species with well characterised endocrinology and physiology (e.g. rats and mice) representative of the animal classes of concern. Notwithstanding this, the ultimate aim would be to minimise the number of species used in such tests, therefore, it is of primary importance to establish end-points which might be extrapolated from species to species. This can only be achieved by cross-validation of effects seen in different phyla/classes using the same chemical and dosing regime together with the same route of exposure.

PRIORITISATION SCREENS *IN VIVO*

Any successful prioritisation screen for EDCs should have a limited number of relevant end-points and be of short duration. Existing regulatory tests, for mammals, of short duration are generally inadequate because they are designed to assess acute(single dose) toxicity of a given chemical and do not employ

endocrine end-points. In addition, the doses employed are extremely high, which can make the interpretation of any data difficult due to overt toxicity. Similarly, there are no non-mammalian regulatory tests which would be appropriate for prioritising chemicals for further testing of potential EDCs in sub-chronic or chronic models. In contrast, there are a number of non-regulatory bioassays, both mammalian and non-mammalian which could be developed for use in identifying EDCs; some of these are described briefly in the following section.

Mammals

The rodent uterotrophic and vaginotrophic models have been used for many decades as bioassays for oestrogen and anti-oestrogen chemicals [93–95]. Due to the diversity in the designs employed, there is presently no definitive and generally accepted protocol for either of these assays. Although mice can in some cases be more sensitive than rats, the rat is the species most widely employed in routine toxicology, thus making it the preferred default species. The general protocol involves administration of the test chemical, either orally or by injection, to ovariectomised animals. Most studies appear to adopt a three day treatment period with sacrifice on the fourth day. Although some authors have indicated that ovariectomy is unnecessary, in order to ensure a consistent response, it is suggested that ovariectomised or immature non-ovariectomised animals are preferable. A variety of end-points are possible for the uterotrophic assay. These include uterine weight, mitotic rate, histopathological structural changes and a range of complex biochemical end-points. In terms of a simple screening tool, it is not necessary to understand the mechanics of any effects seen and therefore, the use of uterine weight alone is probably a sufficient measure of uterine growth. Similarly, in the vaginotrophic model, vaginal washing followed by histological examination of the cells obtained has been suggested as a simple indicator of the induction of oestrous [95] and may therefore be preferable as a first level screening tool.

Prostatic weight was recently proposed as an initial screen for androgenic activity [96]. As is the case in the uterotrophic assay, the use of castrated animals permits greater control of the model and facilitates detection of androgenic activity. The tests generally involve administration of the treatment to young adult males for 15 days or more, followed by subsequent removal of the prostate. As is the case in the uterotrophic assay, weight increases in the prostate may arise as a result of other mechanisms, e.g. tissue oedema or inflammation. This is, however, probably not a significant problem when such a test is used as an initial screen.

There are also a large number of mammalian androgen-responsive genes that are known and for which antibodies and molecular probes are present. For example, rat urinary protein, alpha-2u-globulin is synthesised in the liver in response to androgens and is secreted into the blood. Its gene is controlled by an androgen response element and it would therefore be expected to be sensitive to anti-androgens.

Non-mammalian models

There are a number of non-regulatory protocols which have the potential for use in prioritisation screening in non-mammalian vertebrates. One of the most widely used methods for monitoring oestrogenic activity in oviparous species is the assay for vitellogenin [97], a female yolk protein precursor, that is synthesised by the liver in response to oestrogenic stimulation. In mature female fish, this protein is sequestered by the growing oocytes in the ovary and cleaved to form the major constituents of yolk. Males or immature females have also been demonstrated to produce this protein, which can be detected in the blood plasma, in response to exogenous oestrogenic stimulation. The test chemical can be administered via injection, orally or, in the case of aquatic organisms, via the water. A response should be measurable, at least in fish, in 3–4 days. Although the scientific basis for the selection of this protein as a biomarker for oestrogenic stimulation is well established, the toxicological and ecological significance of the production of this protein is still in question. It is known, however, that excessive amounts of VTG can result in kidney damage [98]. In addition, the production of this protein has been associated with an inhibition of testicular growth [33]. Assays using RIA or ELISA techniques to directly measure vitellogenin are well established and have proved suitable for both field and laboratory experiments on a wide range of chemicals [97,99,100]. The only problems relate to a lack of

standardisation of the study protocols together with the highly species-specific responsiveness of the antibodies used to date. Future research efforts should perhaps be directed towards widely responsive or even universal vitellogenin immunoassays.

Presently, no equivalent marker for androgenic activity is available. It is known, however, that in certain fish species, secondary sexual characteristics are androgen responsive. For example, in the sexually dimorphic poecilid mosquito fish, females are capable of responding to androgenic stimulation by developing male characteristics within a relatively short time of exposure [101]. It should be noted, however, that the factors controlling such responses, are in some cases, species specific. Consequently, it will be necessary to determine the sensitivity and specificity of such models before advocating their use.

In birds, both comb and wattle growth are very sensitive to androgenic and anti-androgenic stimulation [102]. Potential EDCs could be directly applied to the comb of 2–3 day old chicks daily for a period of 7 days after which the chicks are sacrificed and compared with a positive control. In addition, in female birds, the oviduct weight can be used as an indicator of oestrogen exposure and is analogous to the mammalian test of uterine weight.

LIFE CYCLE TESTS

The endocrine system is extremely complex and therefore the selection of end-points for any test is problematical. Every event is a chain of interdependent, mutually controlling, and cyclic reactions. The final result is frequently the product of a whole series of phenomena, all of which must occur in a properly timed sequence if they are to culminate in a normal and predictable effect on the end organ they affect. Thus, in order to understand the intricacies of the hormonal interplays that produce the final result, for example, copulation, one must think of the whole chain of events rather than the single events that make up the chain. For example, gland A secretes hormone A which stimulates end organ 1 to grow and gland B to secrete hormone B. In turn, hormone B modifies the rate of secretion of hormone A from gland A and stimulates end organ 2. Some interrelations in the endocrine system are simpler than this, and most are more complex, but basically they all follow the same scheme. If a chemical is to be adequately tested, therefore, it is dangerous to examine only a single end-point, since almost any process which is involved in reproduction and/or development can be affected by exposure to a single chemical, or mixture. In addition, chemicals which are administered maternally may affect the offspring, rather than the mother. Whole life-cycle or partial life cycle tests are therefore the most preferable method of testing, although they are often more complex. These tests should be used as tools to investigate the reproductive and developmental effects of chemicals that have been selected previously, using *in vitro* or *in vivo* prioritisation screens. In addition, their design would allow the selection of a more modest number of end-points, rather than a single one. These end-points should be quantifiable across a wide dose-range. In addition, the importance of both negative and positive controls cannot be over stressed. Data generated in life cycle tests, whether short-term (partial) or long term(whole) will probably be used in a risk assessment context. Therefore, it is at this stage that close attention should be paid to the method/route and timing of exposure. Testing should be carried out on eggs/embryos or early life stages as well as on mature adults in order to determine any differential effects due to the time of exposure. Indeed, as outlined previously, there has been some suggestion that early life stages may be more sensitive than during adulthood and hence this procedure will also establish whether one or the other (or both) is more appropriate for hazard identification.

Early life stage tests

There is now a growing body of evidence which suggests that exposure to EDCs during early life can affect/disrupt the growth and differentiation of the reproductive tract and accessory organs [103] in all vertebrates, as already outlined above. These types of study would require a minimal amount of experimental effort, as exposure to a test chemical or mixture would only occur during the relatively short period of sexual differentiation after which, the animals could be transferred to a ‘chemical-free’ environment. Fundamental research is also needed to evaluate the consequences of any organisational effects seen, e.g. differentiation of a female reproductive tract in a genetically male animal. Indeed, these

early life stage tests could be extended to a stage when the animals have reached maturity in order that gamete quality and quantity, and/or reproductive success could be evaluated. During the development and validation of such tests, it will also be important to establish a minimum period of exposure in order that the test can be carried out as quickly as possible. To illustrate this point, Kelce & Wilson [104] suggested that dosing chemicals to newly weaned male rats for 1 month would be a useful assay for anti-androgens, causing a reduction in the weight of the accessory sex organs. In a more recent study [105], however, a shorter exposure period of fourteen days was reported to be sufficient, whilst Mylchreest *et al*.[106] report that the period spanning lactation may be more important than the period of organogenesis [106]. Exposure to EDCs may also alter physical growth and/or cellular organisation of the gonadal tissue. For example, in juvenile or adult fish, the growth and development of the gonad is sensitive to endocrine disruption; exposure of male rainbow trout to nonylphenolic chemicals can retard the growth of the testes and inhibit development [33]. It is also important to note that measuring organ weight as an end-point after exposure to a test chemical can produce results which are extremely difficult to interpret. For example, the gonad weight might be expected to increase as a result of exposure to an androgenic substance and therefore the gonadosomatic index (gonad weight relative to body weight) would also increase. Similarly, an inhibition of growth would also cause an increase in the gonadosomatic index, even when there are no differences in the absolute weights of the gonads. In one study, increases in the relative weights of the liver testes and epididymides of the control animals were reported over a 14 day period, concomitantly with a decrease in the relative weights of the kidney, seminal vesicles and prostate. The authors of this study caution that large group sizes will be required, together with in-house control tissue weight data in order to successfully validate such assays [105]. These problems are confounded by the fact that the reproducibility of any rodent-based assay may also depend on the strain of rat used.

In most vertebrate groups, there are currently no internationally agreed regulatory tests that are specifically designed to assess sexual differentiation and development. In some cases, however, existing tests do incorporate such end-points(e.g. nipple retention and anogenital distance), particularly in mammals, where the standard developmental toxicity test is capable of detecting the effects of steroid hormones on structural development. Indeed, this test has very recently been made even more responsive by extension of the dosing period through the end of gestation.

Sub-chronic tests which focus on exposure during early life stages are also needed in non-mammalian vertebrates such as fish. It has already been mentioned that exposure of fish to hormones, their antagonists, or mimics, during sexual differentiation (embryonic or after hatching) [18,19,107,108] can result in disruption of the reproductive tract (could be irreversible) inhibited spermatogenesis, for example, and sterility. The development of an oviduct, for example, in a genetically male fish, has been proposed as an biomarker of oestrogen exposure [18]. The exposure period may only need to be short (a matter of several days or weeks in some species), and the animals could be sampled immediately after the end of the exposure period. Histopathology would normally be required to detect more subtle changes in gonad structure. Knowledge of the period of sexual differentiation in the test fish species is vital for the success of this bioassay, since this may vary vastly from species to species. In addition, the test species should be well characterised with respect to the regulation of sexual differentiation. In some fish species, such as the channel catfish, exposure to androgens causes feminization, whilst in other species, water temperature can also affect sex determination and sex ratios. Some fish change sex (sequential hermaphrodites) as a normal part of their life cycle, whilst others undergo a transient period of intersexuality during the juvenile phase.

In fish, the US EPA has a fish early life stage toxicity test [109] in which exposure is at fertilisation through hatch and first feed. The end-points include successful hatch, survival, growth and developmental/behavioural abnormalities. With few modifications, this test could be used with an appropriate test species. Indeed, Arcand-Hoy & Benson [110] recently proposed the use of the Japanese medaka as a test species both at the early life stage, and during adulthood. This species has the advantage that it is sexually mature six-eight weeks post hatch, whist many larger fish species may take up to two years to reach sexual maturity. Other suitable species with clearly defined reproductive behaviours and short life cycles include guppy, goldfish, sheepshead minnow, fathead minnow, fundulus, and the zebrafish. Some of these species (e.g. goldfish) are suitable because a reasonable amount is known about

their basic endocrinology. Other species, such as the fathead minnow, sheepshead, zebra fish and medaka are appropriate because they are already used in regulatory chronic toxicity studies. Other species are attractive on the basis of understanding of the basis of genetics underlying development and reproduction (e.g. *fundulus*, zebra fish).

Other classes of non-mammalian vertebrate are also sensitive to endocrine disruption during early development. For example, in amphibians reptiles and birds, sexual differentiation and gonadal development are both dependent on hormones, particularly oestradiol, and are therefore sensitive to endocrine disrupters. In birds, sexual differentiation is also dependent on hormones, particularly oestradiol. A suitable test might include the morphological alteration in the differentiation of the reproductive tract of embryonic or recently hatched animals. In reptiles and birds, topical application of the test chemical is suggested as a more preferable and realistic route of exposure than injection; the latter, although widely used, is considered to be completely unnatural and should not be used for anything other than prioritising chemicals for further testing. The process of sexual differentiation in all reptiles is, to a large extent, temperature dependent, although it can be affected by exposure to anti-hormones or hormones. Consequently, reptiles may not be the most preferable type of animal to use routinely in a regulatory screen. Conversely, amphibians could be a very useful group of animals for research into endocrine disruption. They are primarily terrestrial, although they are dependent on water for reproduction. They are also noted for their diverse reproductive strategies, including external fertilisation, internal fertilisation, oviparity, ovoviviparity and viviparity. The majority lay their eggs or young in water, and therefore it should be relatively easy to expose them to any test chemicals via this route. Indeed, so much is known about embryological development and metamorphosis in, for example, *Xenopus laevis*, that it has great potential as a test species for endocrine disruption. Sex reversal of tadpoles has been accomplished with genetically male *Xenopus* [111]exposed to oestrogens, and Ramsdell *et al.* [112] obtained the same response with nonylphenol. Several laboratories will be working with this test in the near future, in order to validate it. Thyroxine controls development and metamorphosis in amphibians and therefore this species is of potential use in the study of thyroid hormone modulators (some PCBs, for example). Furthermore, short term bioassays, in which embryo development is assessed in *Xenopus laevis*, already exist, and are commercially available. For example, the FETAX 96-h embryo toxicity test was designed to assess the teratogenic potential of test chemicals [45]. The outcome of this assay is assessed by the concentration required to induce 50% mortality (96-h LC50), the concentration required to deform 50% of the individuals (96-h EC50) and the minimum concentration required to inhibit growth (MCIG). From these variables, a teratogenic index is calculated by dividing the LC50 by the EC50. The FETAX test is well suited for testing either pure compounds or complex mixtures, such as industrial effluents or mixtures found at hazardous waste sites. Indeed, with little modification (allowing the individual tadpoles to undergo metamorphosis in clean water, after exposure), this test has been successfully used to evaluate the developmental toxicity of some retinoic acids and, perhaps more importantly, the metabolites and the photoisomers of the pesticide s-methoprene [32] that are retinoic acid mimics. This recent finding is of great importance as it serves to illustrate further that toxicity testing for new and existing chemicals must be extended to their metabolites and degradation products. Indeed, the environmental hazard posed by s-methoprene could have been detected using existing tests and procedures prior to its use, if only a full examination of its natural degradation and metabolism had been carried out.

Adult life stage tests

Sub-chronic tests in which reproductive capacity is assessed could also be carried out on adult animals. End-points to be evaluated during the exposure period could include mating behaviour, secondary sexual characteristics(e.g. colour in fish, pattern formation in amphibians), time to first hatch, fecundity(number of eggs produced per female) or number of offspring (in non-egg layers), sperm production, sex steroid concentrations, and vitellogenin induction in male oviparous vertebrates. In amphibians, new tests would have to be devised to evaluate thyroid function.

Plasma steroid concentrations are usually measured using RIA or ELISA techniques with monoclonal antibodies to mammalian oestradiol and testosterone. In some fish, 11-ketotestosterone may be a more

proximal measure of androgenic activity than testosterone [113]. Antisera to this hormone can, however, be difficult to obtain. Similarly, antisera for the maturation-inducing progestrogen steroids are limited in availability. As with vitellogenin, steroids are more difficult to measure in small animal species because of a limited blood volume; however steroid measurements of extracts of whole body or gonadal tissue may be useful as alternative indicators of steroidogenic activity. Another concern is that fluctuations normally occur through the course of the reproductive cycle and, in some species, throughout the day. Data on these types of cyclical variations are not available in many species.

The ratio of gonad weight to body weight is a useful screening end-point for indication of (anti-) oestrogenic/androgenic effects because it is easily measured and applicable to both sexes. Changes in this parameter are indicative of changes in circulating hormone levels and even likely reproductive success. Its major shortfall relates to the fact that it changes rapidly during the reproductive cycle of annual breeders; comparisons can therefore only be made in animals which are at the same stage of gameteogenesis.

The mating behaviour of any animal which does not have a particular, easily recognisable sexual pairing behaviour might not be an advisable parameter to attempt to measure, as it would be difficult to quantify. In addition, some secondary sexual characteristics, such as coloration, are highly subjective and variable. The evaluation of gamete quality is a highly integrative end-point that reflects a variety of mechanisms affected by EDCs and is generally applicable to all species. Gamete size, ovulation, and hydration can all be assessed with relative ease. Furthermore, the development of methods to assess sperm production in some vertebrate groups (e.g. fish) is needed. This may include sperm counting with a haemocytometer, the use of Coulter counter cells, image analysis with a computer assisted sperm tracker [114], or a more general assessment of sperm production, using the spermatocrit [115]. There are many difficulties involved in assessing sperm quality however; the only definitive method is to examine fertilisation success and this is to be encouraged whenever possible.

WHOLE LIFE CYCLE TESTS

The highest degree of complexity for testing the effects of chemical substances on development and reproduction is manifested in whole life cycle or multigeneration studies. It is in this area that current testing methodologies are the most developed, particularly in mammals where the current OECD guidelines for laboratory mammals evaluate many end-points that are under endocrine control. Most pertinent of these for assessing effects on reproduction and development are the one and two generation tests. These chronic studies do include exposure during the perinatal period and the designs are very detailed with numerous end-points. Criteria evaluated include fertility, fecundity, sexual behaviour, gametogenesis, ovulation and spermiation, implantation, development of specific tissues and organs, maintenance of pregnancy, parturition, growth and viability of offspring, lactation and maternal behaviour. All of these parameters are sensitive to endocrine-mediated toxicity and therefore tests such as these would be appropriate after some modification to include additional end-points such as sperm production, sperm motility, accessory sex organ weight, daily vaginal cytology (in females) to evaluate the duration of the oestrous cycle, gonadal histopathology, and sex steroid production. The two generation reproductive toxicity tests are regarded as the most rigorous of the current tests, since they permit an in-depth study of the growth, development and sexual functionality of the F1 generation as well as the monitoring of the subsequent (F2) generation through to weaning. Examination of offspring (such as the F1 or F2), which have been exposed to chemicals *in-utero*, should include end-points such as the timing of vaginal opening, testicular descent, anogenital distance, nipple retention, accessory sex organ weight and vaginal cytology.

Similarly, in birds, there are existing tests which could be modified to evaluate end-points for endocrine disruption. Specifically, the US EPA has tests under the Federal Insecticide Fungicide and Rodenticide Act (FIFRA), and the OECD using northern bobwhite quail , mallard duck and Japanese quail. The mallard test can evaluate egg-shell thinning, but the other species are not generally susceptible to this end-point. This test requires more than 20 weeks to complete, including acclimatisation and pre-reproductive 8 week exposure followed by 10 weeks of egg collection. Both males and females are exposed via feeding and reproduction is assessed using pairs. Egg viability, hatching success (in

incubators) and survival are all examined. The strengths and weaknesses of this test are discussed in [116]. In general, this test is excellent for chemicals which have a propensity to bioaccumulate or are persistent, however, artificial incubation removes parental behaviour which may be affected by exposure to the test chemical. In addition, it does not serve as an indicator of oestrogenic/androgenic effects and does not evaluate the F1 after hatch. A proposed modified test [117]would employ Japanese quail pairs, rather than the mallard because the Japanese quail has a short generation time and thus makes the test a lot shorter. After only 10 weeks (including the pre-reproductive period), shell thickness, sex ratios of the progeny, gonad weight, sex accessory gland weight, oviduct weight of hatched chicks, and histopathology of the reproductive tract could all be examined. The offspring could be sacrificed at only 8–10 days post hatch since at this stage , the sex cells in the gonads are clearly defined. Key indicators of change in the gonads of the offspring would include polyovular follicles in the ovaries of females and primary follicles in the cortex of the testes of males. This test could easily be extended; the offspring could be raised to adulthood and pair bred to evaluate shell quality, fertility, embryo hatchability and growth.

In fish, as is the case in birds and mammals, existing a whole life cycle toxicity testing guidelines need to be adapted to include procedures which would allow analysis of endocrine-specific effects. In a whole life-cycle test, exposure would be over a two to four week period, whilst the fish are in mated pairs or groups. A wide range of end-points could then be examined, including sexual behaviour, secondary sexual characteristics, time to first hatch, fecundity, sperm production, steroid concentrations and vitellogenin induction. In multigenerational tests, the F1 generation could also be examined with or without exposure to the test chemical/mixture. Sexual differentiation, gonadal growth and development and hormone production and/or the production of gametes are all suggested parameters. In all tests, it is important to determine the significance of changes in hormone or protein synthesis, by linking them to biological end-points such as gonadal growth, for example. Another point ,that is well worthy of a mention here, is that almost all of the methods that are suggested in this manuscript involve sacrificial sampling at the end of the experiment and thus, are destructive. Whilst this presents few problems in laboratory-based studies, these methods are hardly suitable for use in monitoring wild populations for endocrine effects. There is, therefore, a real need for the development of non-destructive biomarkers that are indicative of biological effects. For example, intersexuality in fish, as a result of exposure to oestrogenic contaminants, may be widespread in wild populations of fish; a non-destructive biomarker for this phenomenon would, therefore, be most useful for biomonitoring. It is in this area that the development of molecular biomarkers that are indicative of deleterious biological/physiological effects is required. If, for example, the gonadal tissue of intersexual fish expressed certain novel genes, a small tissue biopsy of a live animal would allow a non-sacrificial method for the assessment of these effects.

INVERTEBRATES

Extrapolation of the effects of any chemical or mixture from vertebrates to invertebrates is almost impossible. This is largely due to the fact that the endocrinology of none of the test species is sufficiently understood to include in any endocrine testing regime. Indeed, endocrine systems in invertebrates are poorly understood; each phylum has unique hormones which are not typical of vertebrates and therefore, fundamental studies are required to assess the applicability of invertebrates in the assessment of EDCs. Invertebrates are, however, essential for the health of the ecosystem and are potentially invaluable for evaluating effects in the field, due to their ubiquitous distribution. In the aquatic environment, the greatest species diversity of invertebrates is in shallow waters which can receive contaminated effluents (tributyl tin example). Furthermore, although their role is uncertain, oestrogenic and. androgenic hormones are found in every invertebrate class examined so far [118]. In echinoderms and gastropods, for example, they control oocyte growth. In arthropods, however, ecdysones are the predominant hormones; structurally similar to vertebrate steroidal hormones, they mediate differentiation, growth and reproduction in nematodes and molluscs. They are also important in vitellogenesis and moulting [119,120] Another well-known hormone, juvenile hormone is known to regulate embryogenesis, development and reproduction [121]. Both field and laboratory studies have shown that these animals are susceptible to EDC effects, e.g. testosterone and tributyltin [48]. Some endocrine disrupting chemicals,

such as nonylphenol, have been tested on various species of invertebrate with variable effects ranging from no effects at all to effects on fecundity, or the development of secondary sexual characteristics. The variability in the effects seen is perhaps not surprising when one considers the differences in the endocrine systems between these groups of animal. The assumption that a chemical, such as nonylphenol, will have similar effects on invertebrates because it has effects on vertebrates is contradictory in this respect. Indeed, recent research on the water flea (Daphnia magna) suggests that some chemicals which are oestrogenic in vertebrates affect moulting, not reproduction [122]. Notwithstanding this, agricultural chemicals which have been specifically developed to disrupt the endocrine system of insects are a potential threat to wild species of invertebrate with similar endocrinology. Indeed, it is in this area that research into EDCs for invertebrates and their effects should be developed as it is entirely possible that these chemicals are present in the environment. The ecdysteroid receptor sequences of several species of crustacean and insect, for example, are now well defined [123] and therefore, research into the possibility that certain pollutants might interact with these receptors could be a good starting point. *In vitro* tests using invertebrate steroid receptors and or expression vectors containing these receptors could be developed and environmentally prevalent chemicals could be tested rapidly and efficiently for their ability to act as receptor agonists or antagonists.

CONCLUSION

In conclusion, existing methods for the assessment of reproductive and developmental effects of endocrine disrupting chemicals on wildlife are not sufficiently validated for use in any regulatory testing regime. Several methods are, however broadly suitable and could be applied, after modification and standardisation. The structural conservation of hormones and their receptors should enable investigators to limit prioritisation screening to a few classes of vertebrates, although a substantial degree of cross validation between species and standardisation of existing end-points is an essential research need. In addition, the development of short-term *in vivo* prioritisation screens to complement *in vitro* and SAR approaches is needed.

Despite the absence of internationally agreed test methods for EDCs, several countries have already taken regulatory action. In Austria, Belgium, France, Germany, Sweden and the UK, for example, expert committees have been set up in order to develop National Action Plans for research into EDCs. The UK made the first move with a strategic paper from the environment agency in January 1998. Italy and Greece entered the planning stage in September 1997. Some member states have already proposed controls on known EDCs. These include the phasing out of phthalates in Denmark, control of the use of nonylphenol ethoxylates (NPEOs) in pesticides(Denmark, Belgium, France, Sweden and Finland), and the withdrawal from the market of babies teething rings that give off phthalates in Denmark, Spain, Italy and Sweden. In the USA, The Food Quality Protection Act and the Safe Drinking Water Act were both passed in 1996 and explicitly extend testing and regulatory authority to endocrine disrupting chemicals in food-grade pesticides and as drinking water contaminants. Their time deadlines are to implement screening and testing by August 1999. To assist the agency in developing a sound flexible approach to EDC testing, a committee of experts representing industry, environmental groups, public health groups, environmental justice groups, academia and government has been established (EDSTAC). Several decisions have already been made including the consideration of human and ecological effects, oestrogen, androgen, anti-oestrogen, anti-androgen, and thyroid effects, of single compounds and mixtures.

REFERENCES

1 Bitman, J., Cecil, H.C., Harris, S.J., Fries, G.F. *Science*, 1969, **162**, 371–372.

2 Dodds, E.C., Lawson, W. *Proc Royal Soc. Lon. B.*, 1938, **125**, 222–232.

3 Kavlock, R.J., Daston, G.P., Derosa, C., Fennercrisp, P., Gray, L.E., Kaattari, S., Lucier, G., Luster, M., Mac, M.J., Maczka, C., Miller, R., Moore, J., Rolland, R., Scott, G., Sheehan, D.M., Sinks, T., Tilson, H.A. *Environ. Health Perspect.*, 1996, **104**, 715–740.

4 Ankley, G., Mihaich, E., Stahl, R., Tillitt, D., Colborn, T., McMaster, S., Miller, R., Bantle, J., Campbell, P., Denslow, N., Dickerson, R., Folmar, L., Fry, M., Giesy, J., Gray, L.E., Guiney, P., Hutchinson, T., Kennedy, S., Kramer, V., LeBlanc, G., Mayes, M., Nimrod, A., Patino, R., Peterson, R., Purdy, R., Ringer, R., Thomas, P., Touart, L., VanderKraak, G., Zacharewski, T. *Environ. Toxicol. Chem.* 1998, **17**, 68–87.

5 Tattersfield, L., Matthiessen, P., Campbell, P., Garandy, N., Lange, R.*SETACEurope/OECD/EC Expert Workshop on Endocrine Modulators and Wildlife: Assessment and Testing*; SETAC- Europe, Brussels, Belgium: VELDHOVEN, The Netherlands, 10–13 April, 1997.

6 Ratcliffe, D.A. *Nature*, 1967, **215**, 208–210.

7 Newton, I. *Environ. Poll.* 1988, **55**, 29–40.

8 European Commission. 'European Workshop on the Impact of Endocrine Disrupting Chemicals on human health and wildlife'1996.

9 *Chemically-Induced Alterations in Sexual and Functional Development:The Wildlife/Human Connection*; Colborn, T., Clement, C., Eds., Princeton Scientific Publishing Company., Inc.: Princeton, New Jersey, ; Vol. 1992, **21**, pp 403.

10 McLachlan, J.A. *Estrogens in the environment II*; Elsevier: New York, 1985.

11 Whitten, P.L., Naftolin, F. *Steroids*, 1992, **57**, 56–61.

12 Stilman, R.J. *Am. J. Obstet. gynecol.*, 1982, **142**, 905–921.

13 Bern, H.A., Mills, K.T., Hatch, D.L., Ostrander, P.L., Iguchi, T. *Cancer Lett.* 1992, **63**, 117–124.

14 Gray, L.E., Ostby, J., Ferrell, J., Rehnberg, G., Linder, R., Cooper, R., Goldman, J., Slott, V., Laskey, J. *Fund. Appl. Toxicol.*, 1989, **12**, 92–108.

15 VomSaal, F., Bronson, F. *Biol. Reprod.*, 1978, **19**, 842–853.

16 Vom Saal, F. *J. Reprod. Fertil.*, 1989b, **86**, 457–471.

17 Piferrer, F., Donaldson, E.M. *Aquaculture*, 1989, **77**, 251–262.

18 Gimeno, S., Gerritsen, A., Bowmer, T., Komen, H. *Nature*, 1996, **384**, 221–222.

19 Gray, M.A., Metcalfe, C.D. *Environ. Toxicol. Chem.*, 1997, **16**, 1082–1086.

20 Ungerer, J., Thomas, P. *Mar. Environ. Res.*, 1996, **42**, 167–171.

21 Adkins-regan, E., Ottinger, M.A., Park, J. *J. Exp. Zool.*, 1995, **271**, 466–470.

22 Greco, T., Furlow, J., Duello, T., Gorski, J. *Endocrinol.* 1991, **129**, 1326–1332.

23 Fishman, R.B., Branham, W.S., Streck, R.D., Sheehan, D.M. *Biol. Reprod.*, 1996, **55**, 1221- 1230.

24 Greco, T., Duello, T., Gorski, J. *Endocr. Rev.*, 1993, **14**, 59–71.

25 Trudeau, V.L., Wade, M.G., Vanderkraak, G., Peter, R.E. *Can. J. Zool.*, 1993, **71**, 1131–1135.

26 Drouin, J.D., Lagace, L., Labrie, F. *Endocrinol.*, 1976, **99**, 1477–1481.

27 Frawley, L.S., Neill, J.D. *Endocrinol.*, 1984, **114**, 659–663.

28 Thomas, P. *Mar. Environ. Res.*, 1989, **28**, 499–503.

29 Cooke, P.S. *Anim. Reprod. Sci.*, 1996, **42**, 333–341.

30 Chambon, P. *Faseb Journal*, 1996, **10**, 940–954.

31 Harmon, M.A., Beohm, M.F., Heyman, R.A., Mangelsdorf, D. *J. Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 6157.

32 La Clair, J.J. L., Bantle, J.A., Dumont, J. *Environ. Sci. Tech.* 1998, **32**, 1453–1461.

33 Jobling, S., Sheahan, D.A., Osborne, J.A., Matthiessen, P., Sumpter, J.P. *Environ. Toxicol. Chem.*, 1995, **15**, 194–202.

34 Harries, J.E., Sheahan, D.A., Jobling, S., Matthiessen, P., Neall, P., Sumpter, J.P., Tylor, T., Zaman, N. *Environ. Toxicol. Chem.* 1997, **16**, 534–542.

35 Blackburn, M.A., Waldock, M.J. *Water Res.* 1995, **29**, 1623–1629.

36 Jobling, S., Nolan, M., Brighty, G., Tyler, C.R., Sumpter, J.P. *Environ. Sci. Tech.* In Press, 1998.

37 Carlsen, E., Giwercman, A., Keiding, N., Skakkebaek, N.E. *B. M. J.* 1992, **305**, 609–613.

38 Bahadur, G., Ling, K.L. E., Katz, M. *Human Reprod.*, 1996, **11**, 2635–2639.

39 Demouzon, J., Spira, A., Thonneau, P., Multigner, L. *B. M. J.*, 1996, **313**, 43–43.

40 ECETOC Document No. 33 'Environmental Oestrogens' A Compendium of Test Methods, ' Avenue E.Van Nieuwenhuyse, 4, B-1160 Brussels, Belgium, 1996.

41 Fry, D.M., Toone, C.K. *Science*, 1981, **213**, 922–924.

42 Fry, D.M., Toone, C.K., Speich, S.M., Peard, R.J. *Stud. Av. Biol.* 1987, **10**, 26–43.

43 OECD guidelines for testing chemicals, guideline 206. Avian reproductive test.

44 OECD guidelines for testing chemicals. Guideline 210. Fish Early Life Stage Test, 1992.

45 Bantle, J.A. In Fundamentals of Applied Toxicology 2nd ed. Eds., G.M. Rand, Ed., Taylor and Francis: Washington DC; pp 207–230, 1993.

46 Gibbs, P.E., Bryan, G.W., Pascoe, P.L., Burt, G.R. *J. M. B. A. U. K.*, 1987, **67**, 507–523.

47 Bryan, G.W., Gibbs, P.E., Burt, G.R., Hummerstone, L.G. *J. M. B. A. U. K.*, 1987, **67**, 525–544.

48 Bettin, C., Oehlmann, J., Stroben, E. *Helgolander Meeresuntersuchungen*, 1996, **50**, 299–317.

49 Ousterhout, J., Struck, R.F., Nelson, J.A. *Biochem. Pharmacol.* 1991, **30**, 2869–2871.

50 Giger, W., Ahel, M., Koch, M., Laubscher, H.U., Schaffner, C., Schneider, J. *Wat. Sci. Tech* 1987, **19**, 449–460.

51 Kelce, W.R., Monosson, E., Gamcsik, M.P., Laws, S.C., Gray, L.E. *Toxicol. Appl. Pharmaco.*, 1994, **126**, 276–285.

52 Jobling, S., Reynolds, T., White, R., Parker, M.G., Sumpter, J.P. *Environ. Health Perspect.* 1995, **103**, 582–587.

53 Soto, A.M., Chung, K.L., Sonnenschein, C. *Environ. Health Perspect.* 1994, **102**, 380–383.

54 Melancon, M.J., Lech, J.J. *Xenobiotica*, 1979, **9**, 317.

55 Ankley, G.T., Bradbury, S.P., Hermens, J., Mekenyan, O., Tollefson, K.-E. In *Society of Environmental Toxicology and Chemistry-Europe/OECD/EC expert workshop on endocrine modulators, Wildlife Assessment and Testing.*, SETAC Europe Press: Brussels, Belgium, 1998; In Press.

56 Routledge, E.J., Sumpter, J.P. *J. Biol. Chemi.* 1997, **272**, 3280–3288.

57 White, R., Jobling, S., Hoare, S.A., Sumpter, J.P., Parker, M.G. *Endocrinol.* 1994, **135**, 175–182.

58 Waller, C.L., Oprea, T.I., Chae, K., Park, H.K., Korach, K.S., Laws, S.C., Wiese, T.E., Kelce, W.R., Gray, L.E. *Chem. Res. Toxicol.* 1996, **9**, 1240–1248.

59 Tong, W.D., Perkins, R., Xing, L., Welsh, W.J., Sheehan, D.M. *Endocrinol.* 1997, **138**, 4022–4025.

60 Waller, C.L., Juma, B.W., Gray, L.E., Kelce, W.R. *Toxicol. Appl. Pharmacol.*, 1996, **137**, 219–227.

61 Kuiper, G., Enmark, E., Peltohuikko, M., Nilsson, S., Gustafsson, J.A. *Proc. Nat. Acad. Sci. U. S. A.*, 1996, **93**, 5925–5930.

62 Zacharewski, T. *Environ. Sci. Tech.*, 1997, **31**, 613–623.

63 Soto, A.M., Lin, T.M., Justicia, H., Silvia, R.M., Sonnenschein, C. In *Chemically Induced Alterations in Sexual and Functional Development: The Wildlife/Human Connection*; T.Colborn and C.Clement, Eds., Princeton Scientific Publishing Company Incorporated: New Jersey, ; Vol. XXI; pp 295–309, 1992.

64 Wiese, T.E., Kral, L.G., Dennis, K.E., Butler, W.B., Brooks, S.C. *In Vitro Cell Dev. Biol.*, 28A, 595–602, 1992.

65 VanDerBurg, B., Groot, R.P. D., Isbrucker, L., Kruijer, W., DeLaat, S.W.J. *J. Steroid Biochem. Mol. Biol.*, 1992, **43**, 111–115.

66 Soto, A.M., Sonnenschein, C., Chung, K.L., Fernandez, M.F., Olea, N., Serrano, F.O. *Environ. Health Perspect.*, 1995, **103**, 113–122.

67 Abraham, E.J., Frawley, L.S. *Life Sci.*, 1997, **60**, 1457–1465.

68 Jobling, S., Sumpter, J.P. *Aquat. Toxicol.*, 1993, **27**, 361–372.

69 Steinmetz, R., Brown, N.G., Allen, D.L., Bigsby, R.M., BenJonathan, N. *Endocrinol.*, 1997, **138**, 1780–1786.

70 Katzenellenbogen, B.S., Montano, M.M., Ekena, K., Herman, M.E., McInerney, E.M. *Breast Cancer Res. Treat.*, 1997, **44**, 23–38.

71 Benhamou, B., Garcia, T., Lerouge, T., Vergezac, A., Gofflo, D., Bigogne, C., Chambon, P., Gronemeyer, H. *Science*, 1992, **255**, 206–209.

72 Cavailles, V., Garcia, M., Rochefort, H. *Mol. Endocrinol.* 1989, **3**, 552–558.

73 Mader, S., Leroy, P., Chen, J.Y., Chambon, P. *J. Biol. Chem.* 1993, **268**, 591–600.

74 Keller, H., Givel, F., Perroud, M., Wahli, W. *Mol. Endocrinol.*, 1995, **9**, 794–804.

75 Zacharewski, T.R., Berhane, K., Gillesby, B.E., Burnison, B.K. *Environ. Sci. Tech.*, 1995, **29**, 2140–2146.

76 Webb, P., Lopez, G.N., Greene, G.L., Baxter, J.D., Kushner, P.J. *Mol. Endocrinol.*, 1992, **6**, 157–167.

77 Purvis, I.J., Chotai, D., Dykes, C.W., Lubahn, D.B., French, F.S., Wilson, E.M., Hobden, A.N. *Gene*, 1991, **106**, 35–42.

78 Oerterklein, K., Baron, J., Colli, M.J., McDonnell, D.P., Cutler, G.B. *J. Clin. Invest.*, 2475–2480, 1994.

79 Routledge, E.J., Sumpter, J.P. *Environ. Toxicol. Chem.*, 1996, **15**, 241–248.

80 Coldham, N.G., Dave, M., Sivapathasundaram, S., McDonnell, D.P., Connor, C., Sauer, M.J. *Environ. Health Perspect.*, 1997, **105**, 734–742.

81 Shiau, S.P., Glasebrook, A., Hardikar, S.D., Yang, N.N., Hershberger, C.L. *Gene*, 1996, **179**, 205–210.

82 White, R., Jobling, S., Hoare, S.A., Sumpter, J.P., Parker, M.G. *Endocrinol.* 1994, **135**, 175–182.

83 Bicknell, R.J., Herbison, A.E., Sumpter, J.P. *J. Steroid Biochem. Mol. Biol.*, 1995, **54**, 7–9.

84 Harris, C.A., Henttu, P., Parker, M.G., Sumpter, J.P. *Environ. Health Perspect.*, 1997, **105**, 802- 811.

85 Meek, M.D., Clemons, J., Wu, Z.F., Zacherewski, T.B. In *Proceedings of the17th Annual SETAC Meeting*; Washington DC, 1996; pp 443, November 1996.

86 Ashby, J., Harris, C.A., Lefevre, P.A., Routledge, E.J., Sumpter, J.P. *Nature*, 1997, **385**, 494.

87 Arnold, S.F., Klotz, D.M., Vonier, P.M., Collins, B.M., McLachlan, J.A. *Mol. Cell. Endocrinol.*, 1996, **123**, 119–122.

88 Safe, S.H. *Pharmacol. Ther.* 1995, **67**, 247–281.

89 Kelce, W.R., Stone, C.R., Laws, S.C., Gray, L.E., Kemppainen, J.A., Wilson, E.M. *Nature*, 1995, **375**, 581–585.

90 Lundholm, C.E. *Comp. Biochem. Physiol. C-Pharmacol. Toxicol. Endocrinol.*, 1997, **118**, 113- 128.

91 Wallace, K.B. *Environ, Toxicol. Chem.*, 1989, **8**, 1049–1055.

92 Ronis, M.J. J., Walker, C.H. *Comp. Biochem. Physiol. C.*, 1985, **82**, 445–449.

93 Allen, E., Doisy, E.A. *Am. J. Physiol*, 1924, **69**, 577–588.

94 Jones, R.C., Edgren, R.A. *Fertil. Steril.*, 1973, **24**, 284–291.

95 Oconnor, J., Cook, J., Craven, S., Vanpelt, C., Obourn, J. *Fund. Appl. Toxicol.*, 1996, **33**, 182- 195.

96 Cook, J.C., Kaplan, A.M., Davis, L.G., Oconnor, J.C. *Reg. Toxicol. Pharmacol.*, 1997, **26**, 60- 68.

97 Sumpter, J.P., Jobling, S. *Environ. Health Perspect.*, 1995, **103**, 173–178.

98 Herman, R.L., Kincaid, H.L. *Aquaculture*, 1988, **72**, 165–172.

99 Tyler, C.R., Vandereerden, B., Jobling, S., Panter, G., Sumpter, J.P. *J. Comp. Physiol. B- Biochem. Sys. Environ. Physiol.*, 1996, **166**, 418–426.

100 Bon, E., Barbe, U., Rodriguez, J.N., Cuisset, B., Pelissero, C., Sumpter, J.P., LeMenn, F. *Comp. Biochem. Physiol. B-Biochem. Mol. Biol.*, 1997, **117**, 75–84.

101 Howell, W.M., Denton, T.E. *Environ. Biol. Fish.*, 1989, **24**, 43–51.

102 Hennessey, V.J., Galss, R.A., Barnes, S., Vigersky, A.R. *Proc. Soc. Exper. Biol. Med.*, 1986, **182**, 443–447.

103 Gray, L.E., Ostby, J., Wolf, C., Lambright, C., Kelce, W. *Environ. Toxicol. Chem.*, 1998, **17**, 109- 118.

104 Kelce, W.R., Wilson, E.M. *J. Mol. Med.*, 1997, **75**, 198–207.

105 Ashby, J., Lefevre, P.A. *Reg. Toxicol. Pharmacol.*, 1997, **26**, 330–337.

106 Mylchreest, E., Cattley, R.C., Foster, P.M. D. *Toxicol. Sci.*, 1998, **43**, 47–60.

107 Yamamoto, T., Masuda, N. *Gen. Comp. Endocrinol.*, 1963, **3**, 101–110.

108 Piferrer, F., Donaldson, E.M. *Aquaculture*, 1992, **106**, 183–193.

109 Ecological effects test guidelines. Office of Prevention, Pesticides and Toxic substances. Fish early life-stage toxicity test. EPA, 1996.

110 Arcand-Hoy, L.D., Benson, W.H. *Environ. Toxicol. Chem.*, 1998, **17**, 49–57.

111 Villalpando, I., Merchant-Larios, H. *Int. J. Dev. Biol.*, 1994, **34**, 282–285.

112 Ramsdell, H.S., Blandin, D.A., Schmechel, T.R. In *17th Annual Meeting of the Society of Environ. Toxicol. Chem.*, Washington D.C., USA, 1996.

113 Idler, D.R., Bitners, I.I., Schmidt, P.J. *Can. J. Biochem. Physiol.*, 1961, **39**, 1737–1742.

114 Kime, D.E., Ebrahimi, M., Nysten, K., Roelants, I., Rurangwa, E., Moore, H.D. M., Ollevier, F. *Aquat. Toxicol.*, 1996, **36**, 223–237.

115 Bouck, G.R., Jacobson, J. *Trans. Am. Fish Soc.*, 1976, **105**, 534–535.

116 Bennett, R.S., Ganio, L.M. 'Overview of methods for evaluating persticides on reproduction in birds' U.S. Environmental Protection Agency, Corvallis, OR, U.S.A., 1991.

117 Benson, W.H., VanDerKraak, G., Tyler, C., Brugger, K.E., Daston, G., Fry, M., Gimeno, S., Hunger, F., Kolossa, M., Lange, R., Matthiessen, P. In *SETAC-Europe/OECD/EC Expert workshop on endocrine modulators and wildlife: Assessment and testing EMWAT*; L.Tattersfield, P.Matthiessen, P.Campbell, N.Grandy and R.Lange, Eds., SETAC-Europe: Brussels, Belgium, ; pp 59–76, 1997.

118 LeBlanc, G.A. In *Environmental Endocrine Disruptors: An Evolutionary Perspective*; J.LJ Guillette, Ed., Taylor and Francis: London, U.K. In Press, 1998.

119 Charniaux-Cotton, H., Payen, G. In *Endocrinology of Selected Invertebrate Types*; H.Laufer and G.Downer, Eds., Alan R. Liss: New York NY, USA, pp 279–303, 1988.

120 Chang, E.S., O'Connor, D. In *Endocrinology of Selected Invertebrate Types*; H.Laufer and G.Downer, Eds., Alan R.Liss: New York NY, USA, pp 259–278, 1988.

121 Wigglesworth, V.B. *Insect Hormones*; W.H.Freeman: San Francisco, CA, USA, 1970.

122 Zou, E., Fingerman, M. *Bull. Environ. Contam. Toxicol.*, 1997, **58**, 596–602.

123 Chung, A.C. K., Durica, D.S., Clifton, S.W., Roe, B.A., Hopkins, P.M. *Mol. Cell. Endocrinol*., 1998, **139**, 209–227.