CHAPTER 19

Immunotoxicity

MARYJANE K. SELGRADE

19.1 INTRODUCTION

A properly functioning immune system is essential to good health. It defends the body against infectious agents and in some cases tumor cells. Individuals with immune deficiencies resulting from genetic defects, diseases (e.g., AIDS, leukemia), or drug therapies are more susceptible to infections and certain types of cancer, the consequences of which can be life-threatening. On the other hand, the immune system may react to foreign substances that would otherwise be relatively innocuous, such as certain chemicals, pollens, and house dust. The resulting allergic reactions can produce an array of pathologies, ranging from skin rashes and rhinitis to more life-threatening asthmatic and anaphylactic reactions. A crucial part of immune function is the ability to distinguish endogenous components (*self*) from potentially harmful exogenous components (*non-self*). Failure to make this distinction results in autoimmune disease.

Immunotoxicology is the study of undesired effects resulting from the interactions of xenobiotics with the immune system (Figure 19.1). There is evidence that some xenobiotics can cause immune suppression. Xenobiotics can also interact with the immune system to either cause or exacerbate allergic disease. Finally there is growing concern that xenobiotics could have some involvement in autoimmune disease. This chapter provides a brief overview of the immune system, chemicals associated with immune suppression and immune pathologies, and approaches to testing for these effects.

19.2 THE IMMUNE SYSTEM

Cells of the immune system include several types of leukocytes (white blood cells) (Table 19.1), which are derived from bone marrow. T lymphocytes, a subset of immune

Disclaimer: This chapter has been reviewed by the National Health and Environmental Effects Research Laboratory, US Environmental Protection Agency and approved for publication. Approval does not signify that the contents necessarily reflects the views and policies of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

A Textbook of Modern Toxicology, Third Edition, edited by Ernest Hodgson
ISBN 0-471-26508-X

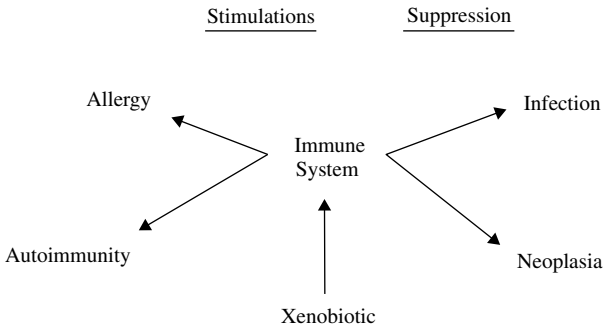

Figure 19.1 Potential consequences of immunotoxicity.

Table 19.1 Leukocytes

Granulocytes (polymorphonuclear leukocytes)
Neutrophils
Eosinophils
Basophils/mast cells[a]
Monocytes
Lymphocytes
Monocytes/macrophages[a]
Natural killer cells

[a]Found in blood/more activated form found in tissues.

cells, undergo differentiation and maturation in the thymus. Leukocytes circulate throughout the body in blood and lymph and populate other lymphoid tissues including the spleen, lymph nodes (scattered throughout the body), tonsils, and adenoids, as well as aggregates of lymphoid tissue in the lung, gut, and skin, which are referred to as bronchus-, gut- and skin-associated lymphoid tissue (BALT, GALT, and SALT). Also immune cells can be recruited to almost any tissue in the body where there is injury or infection. Accumulation of leukocytes in tissues in response to injury is known as inflammation. Cytokines (e.g., interleukins, interferons, and chemokines), soluble mediators produced by immune cells as well as cells outside the immune system, control the maturation, differentiation, and mobilization of immune cells. Immune responses are divided into innate responses directed nonspecifically against foreign substances, and acquired responses directed against specific antigens. There is considerable interaction between these two types of immunity.

Innate immunity provides a rapid, although usually incomplete, antimicrobial defense. Granulocytes, natural killer cells, and macrophages are important mediators of innate immunity. Granulocytes have the capacity to phagocytize (engulf) infectious agents or other types of particles and to destroy or remove them from the tissue. They release a variety of soluble mediators that can kill invading organisms, increase vascular permeability, and recruit more leukocytes to the tissue. Natural killer cells are large granular lymphocytes that nonspecifically kill tumor and virus-infected cells. Macrophages are also phagocytic, can release chemotactic and cytotoxic cytokines, and, when activated, can kill tumor or virus-infected cells. Mediators released from

all of these cells during the acute inflammatory response influence the development of acquired immune responses.

Acquired immunity *specifically* recognizes foreign substances (called antigens) and *selectively* eliminates them. On re-encountering the same antigen there is an enhanced response providing protection against reinfection. Vaccination against infectious agents is based on this principle. T lymphocytes and B lymphocytes (T cells and B cells) are the major players in acquired immunity (Figure 19.2). In both cases there are millions of different clones, groups of immune cells that have specific receptors for a particular antigen. When a cell encounters that specific antigen, clonal expansion occurs; that is, B and T cells with that particular specificity divide and differentiate and are thus activated to respond to the current crisis (e.g., infection). Memory cells develop that represent an enlarged clone of long-lived cells that are committed to respond rapidly, by clonal expansion, upon re-exposure to the same antigen.

B cells recognize native or denatured forms of proteins or carbohydrates in soluble, particulate, or cell-bound form. Activated B cells differentiate into plasma cells and produce antibodies, soluble proteins known as immunoglobulins (Ig), that circulate freely and react specifically with the invoking antigen. There are several classes (called isotypes) of Ig molecules—IgM, IgG, IgA, IgE, and IgD. IgM is the predominant antibody in the primary immune response (following initial exposure to an antigen). IgG usually appears later, following a primary infection, but is the predominant antibody

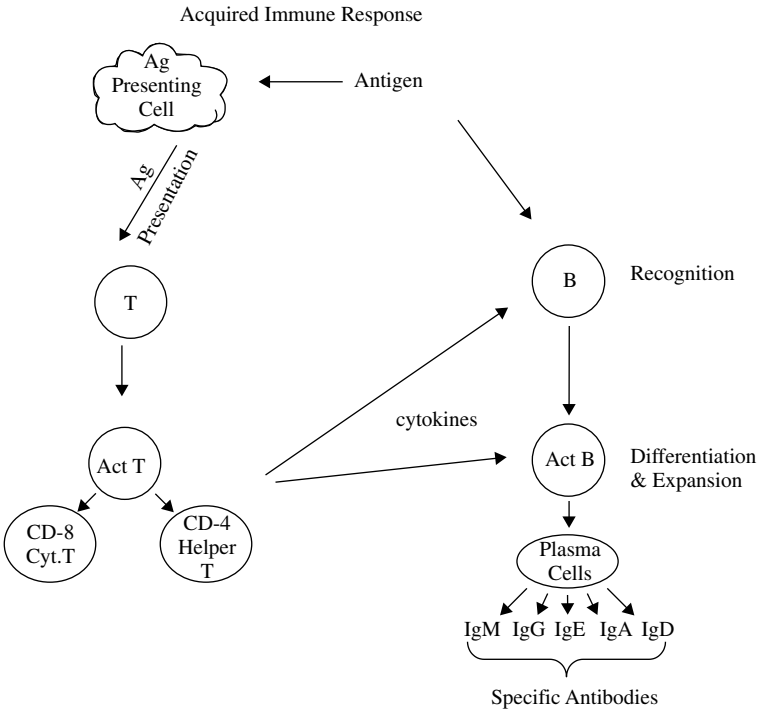

Figure 19.2 The acquired immune response. In response to a specific antigen there is clonal expansion of B cells and subsequent production of antibodies (Ig) specific for that antigen. Antigen presenting cells process and present antigen to T cells. Again there is clonal expansion of cells specific for that antigen.

in the response to subsequent exposures. IgE acts as a mediator of allergy and parasitic immunity. IgA is found in secretions such as mucous, tears, saliva, and milk, as well as serum, and acts locally to block entrance of pathogens through mucous membranes. IgD is mainly membrane bound on B cells. Little is known about the function of this isotype. It does not appear to have a unique role that affects host immunity.

A given B cell will form antibody against just one single antigen; however, during the lifetime of this cell, it can switch to make a different class of antibody. Isotype switching is mediated by T helper cells. B cells recognize two types of antigen: T-independent antigens, which activate the cell without T cell help (predominantly an IgM response), and T-dependent antigens, which required T cell help in order to activate B cells. Most antigens belong to this latter category. Antibodies that specifically recognize microbial antigens can, in combination with plasma proteins known as complement, lyse bacterial cells or neutralize virus. Also microbes complexed with antibody are more readily phagocytized.

T cells recognize antigen that is presented via an antigen-presenting cell (APC) such as macrophages or dendritic cells. APCs process and present short peptide fragments complexed with major histocompatibility (MHC) molecules on the surface of the APC. This processing and presentation is required for T cell activation. There are two major divisions of T cells that are distinguished by expression of different cell surface markers (CD4 and CD8). CD-4 cells are also know as T-helper cells because they provide help for B cell activation. CD-8 cells are also known as cytotoxic T cells because they lyse cells expressing specific viral or tumor antigens.

As indicated above the thymus plays a key role in T cell differentiation. Pre-T cells migrate from the bone marrow to the thymus. As relatively immature cells, T cells express both CD4 and CD8 molecules. As maturation progresses these cells undergo both positive and negative selection. During positive selection only cells that bind to MHC with a certain affinity survive. As a result of this process T cells become MHC restricted; that is, they will only respond to antigen presented in association with MHC. Cells that survive positive selection are potentially able to respond to self proteins. However, before T cells leave the thymus negative selection occurs during which self-reactive cells are removed or functionally inactivated. During the course of positive and negative selection CD4+ CD8+ cells down-regulate the expression of one of these molecules such that mature T cells express only CD4 or CD8. Mature T cells leave the thymus and populate secondary lymphoid organs.

19.3 IMMUNE SUPPRESSION

Experimental studies in laboratory rodents have demonstrated that a diverse array of chemical exposures suppress immune function (Table 19.2). In addition a limited number of clinical and epidemiologic studies have reported suppression of immune function and/or increased frequency of infectious and/or neoplastic disease following exposure of humans to some of these agents. From the description above it is clear there are a number of cellular and molecular targets for chemicals that act as immunosuppressants. Clearly, a chemical that disrupts cell proliferation would affect clonal expansion. Disruption of T cell maturation in the thymus is another potential mechanism for immune suppression. Chemicals may also interfere with receptor ligand binding at the cell

Table 19.2 Selected Examples of Immunosuppressive Agents

Drugs
Cyclosporin A, cyclophosphamide, glucocorticoids (Dexamethazone), azothioprine
Metals
Lead, cadmium, methylmercury, organotins[a]
Pesticides
Chlorodane[a], DDT[a], Dieldrin[a]
Industrial compounds
2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD), polychlorinated and polybrominated biphenyls (PCBs and PBBs), benzene, poly aromatic hydrocarbons[a]
Addictive substances
Cocaine, ethanol, opiates, cannabinoids, nicotine
Air pollutants
Environmental tobacco smoke, ozone, nitrogen dioxide
Microbial toxins
Aflatoxin,[b] ochratoxin A,[b] trichothecenes T-2 toxin[b]
Radiation
Ionizing, UV
Other
Asbestos, diethylstilbestrol (DES), dimethylnitrosamine

[a]Effects in humans are unknown; for all other compound without superscripts changes have been demonstrated in both rodents and humans.
[b]Effects in humans unknown, but veterinary clinicians have noted immunosuppression in livestock ingesting mycotoxins at levels below those that cause overt toxicity.

surface and/or the cascade of signals that lead to transcription of genes responsible for generating and regulating the appropriate immune responses.

Because of the complexity of the immune system, tiered approaches to testing chemicals for immunosuppressive potential have been developed. Like other types of toxicity testing, the first level of the tier (Table 19.3) frequently relies solely on structural end points, including changes in the weight of thymus and other lymphoid organs, histopathology of these organs, or differential blood cell counts. This type of evaluation is convenient because it can be carried out along with an evaluation for other organ systems during routine toxicity testing using one set of animals. However, although these nonfunctional endpoints may be effective in identifying gross (high dose) immunotoxic effects, they are not very accurate in predicting changes in immune function or alterations in susceptibility to challenge with infectious agents or tumor cells at lower chemical doses. Hence the first testing tier (Table 19.3) often includes functional end points designed to assess (1) antibody-mediated responses, (2) T-cell-mediated responses, and (3) NK cell activity. The most commonly used immune function assay in laboratory animals assesses the ability of a mouse or rat to respond to challenge with an antigen, usually sheep red blood cells (SRBC) (Figure 19.3). The response is assessed by determining the number of antigen specific antibody (IgM)

Table 19.3 Tier I Tests (Screen) for Immune Suppression Using Laboratory Rodents

Immunopathology	Hematology: Complete blood count and differential Weights: body, spleen, thymus Histology: Spleen, thymus, lymph node
Antibody-mediated immunity	IgM plaque-forming cell (PFC) response to T cell-dependent antigen (e.g., SRBC)
Cell-mediated immunity	Lymphoproliferative response: T cell mitogens (Con A and PHA) Allogeneic mixed leukocyte response (MLR)
Nonspecific immunity	Natural killer (NK) cell activity

Note: For details on specific assays see M. I. Luster et al., *Fund Appl. Tox.* **10**: 2–19, 1988.

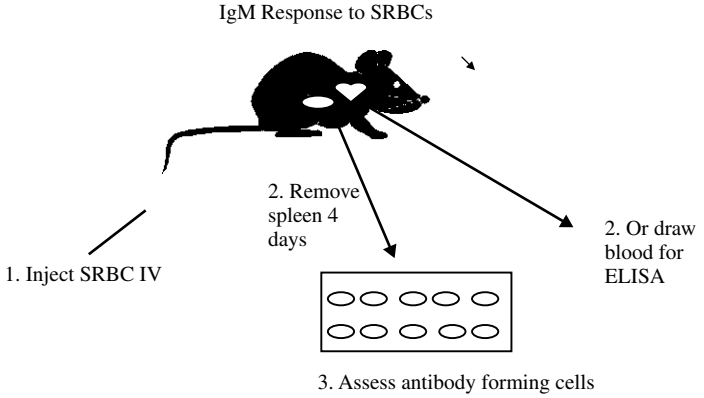

Figure 19.3 Assessing chemicals for immunosuppressive effects. The most common approach to accomplish this goal is to inject chemical and vehicle treated mice or rats with antigen and assess the antibody response. Most often the antigen injected is sheep red blood cells (SRBC); four days later slides are made with a single cell suspension of spleen cells, sheep red blood cells, and complement immobilized in agar. Slides are incubated and spleen cells making antibody against SRBC lyse the surrounding RBCs generating plaques. Plaques are counted to determine the number of antibody forming cells. Alternatively, serum can be obtained and an ELISA assay performed to detect SRBC specific antibody.

forming cells (AFC) in the spleen (Jerne assay) or by assessing antigen specific antibodies in serum using an enzyme-linked immunosorbent assay (ELISA). Because the SRBC is a T-dependent antigen, T and B cells, as well as antigen presenting cells, must be functional to have a successful immunization. Suppression of this response is highly predictive of suppression of other immune function tests and also correlates well with tests that assess resistance to challenge with an infectious agent or tumor cells. The disadvantage to this test is that it usually requires a dedicated set of animals because of the antigen challenge. The most common approach has been to treat the animals for 14 to 28 days with the xenobiotic of interest, inject the antigen at the end of that exposure, and collect spleen or serum 4 to 5 days later. Unlike the tests for antibody-mediated immunity, tier 1 tests for cell-mediated immunity, and natural killer cell activity can be done ex vivo and do not require a dedicated set of animals. However, these tests focus on one cell type and are not as predictive of overall immunocompetence as the antibody assays.

Table 19.4 Tier II More Indepth Evaluation of Immunosuppressive Chemicals

Immunopathology	Quantitation of B and T cell numbers using flow cytometry
Antibody-mediated immunity	IgG PFC to SRBC
	IgM PFC to T cell-independent antigen (e.g., TNP-LPS)
Cell-mediated immunity	Cytotoxic T lymphocyte (CTL) cytolysis
	Delayed hypersensitivity response (DHR)
Nonspecific immunity	Macrophage: phagocytosis, bactericidal/tumoricidal activity)
	Neutrophil: function (phagocytosis and bactericidal activity)
Host resistance models	Response to challenge with infectious agent or tumor cells

Note: For details on specific assays see M. I. Luster et al., *Fund Appl. Tox.* **10**: 2–19, 1988.

When immunosuppressive effects are noted in tier 1, an in-depth evaluation using more sophisticated tests may be carried out (tier 2, Table 19.4). This might include enumeration of lymphocyte subsets (B cells, total T cells, and CD4+ and CD8+) using flow cytometry or assessment of the IgM response to a T-independent antigen in an effort to determine what portion of the immune response is the actual target. Unlike tier 1, tests of cell-mediated immunity in tier 2 require administration of an antigen and subsequent test for cytotoxic T cells (e.g., against an immunizing tumor cell) or a delayed type hypersensitivity response (similar to the response to a tuberculin test). In order to understand the mechanism's underlying immune suppression, a host of other tests can be carried out, including expression of an assortment of cytokines.

Tier 2 also include host resistance models, tests in which an animal is exposed to a xenobiotic and then challenged with an infectious agent or tumor cells. This is considered the ultimate test for an adverse effect on the immune system. However, it should be noted that the amount of immune suppression that can be tolerated is greatly dependent on the dose and virulence of the challenging agent, as well as the genetics of the host. Manipulation of these variables can affect greatly results obtained in host resistance tests.

As in animal studies, human clinical data obtained from routine hematology (differential cell counts) and clinical chemistry (serum immunoglobulin levels) may provide general information on the status of the immune system in humans. However, as with the animal studies, these may not be as sensitive nor as informative as assays that target specific components of the immune system and/or assess function. The assessment of certain lymphocyte surface antigens has been successfully used in the clinic to detect and monitor the progression or regression of leukemias, lymphomas, and HIV infections, all diseases associated with severe immunosuppression. However, there is considerable variability in the "normal" human population, such that the clinical significance of slight to moderate quantitative changes in the numbers of immune cell populations is difficult to interpret. There is consensus within the immunotoxicology community that tests that measure the response to an actual antigen challenge are likely to be more reliable predictors of immunotoxicity than flow cytometric assays for cell surface markers because the latter generally only assesses the state of the immune system at rest. For ethical reasons it is not possible to immunize humans with SRBC. One approach under consideration is assessing responses to vaccines in chemically exposed populations. This approach has been used successfully to demonstrate a link between mild, stress-induced suppression of the antibody response to influenza vaccine and enhanced risk of infectious disease.

There is some debate over how to interpret immunotoxicity data with respect to adversity. The most conservative interpretation is that any significant suppression of an immune response is adverse because a linear relationship between immune suppression and susceptibility is assumed. Supporting this notion is the fact that apparently immunocompetent individuals suffer from infections, suggesting that adverse effects can occur even when known immune suppression is zero. Others argue that there is clearly redundancy and reserve capacity in the immune response and that some suppression should be tolerable. It is impossible to establish a quantitative relationship between immune suppression and increased risk of infection because both the genetics of the host and the virulence and dose of the infectious agent will influence this relationship. Immunocompetence in a population can probably be represented as a bell-shaped curve, such that a portion of the population is highly susceptible to infection, a portion is highly resistant, and the remaining population falls somewhere in between (Figure 19.4). Genetics, age, and preexisting disease all contribute to the risk represented by this curve. In addition the portion of the population at risk is determined by the dose and virulence of any infectious agent that might be encountered. The higher the dose and the virulence, the more people are at risk. Exposure to an immunosuppresive agent shifts the whole bell-shaped curve to the left, thus increasing the population at risk. Unfortunately, it is difficult to determine more quantitatively the relationship between small decrements in immune responsiveness and the degree of change in the population at risk.

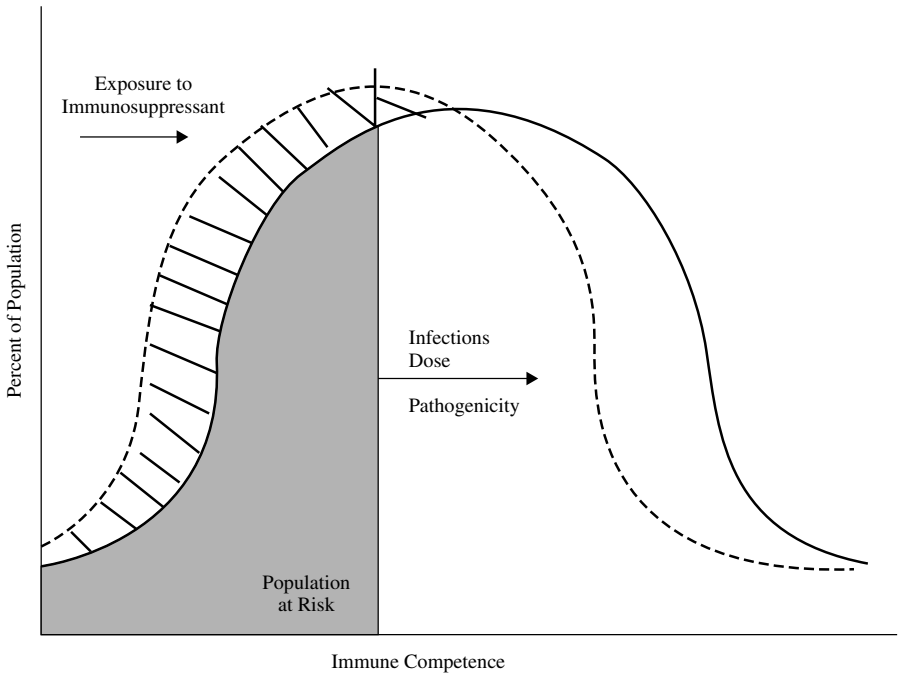

Figure 19.4 Adverse effect of immune suppression. Immune competence is represented by a bell-shape curve. The shaded area represents the population at risk of infection, which increases or decreases depending on the dose and virulence of infectious agents that are encountered. Exposure to an immunosuppressant shifts the whole curve to the left, such that a larger population is at risk for any given infectious challenge.

19.4 CLASSIFICATION OF IMMUNE-MEDIATED INJURY (HYPERSENSITIVITY)

Under certain circumstances immune responses can produce tissue damage. These deleterious reactions are collectively known as hypersensitivity or allergy. Hypersensitivity reactions have been divided into four types (originally proposed by Gell and Coombs) based on mechanism (Table 19.5). In all cases the adverse effects of hypersensitivity develop in two stages: (1) Induction (sensitization) requires a sufficient or cumulative exposure dose of the sensitizing agent to induce immune responses that cause no obvious symptoms. (2) Elicitation occurs in sensitized individuals upon subsequent exposure to the antigen and results in adverse antigen-specific responses that include inflammation.

Type I hypersensitivity (sometimes referred to as atopy) is mediated by an antigen-specific cytophilic antibody (usually IgE) that binds to mast cells and basophils. On subsequent exposure, the allergen binds to these cell-bound antibodies and cross-links IgE molecules, causing the release of mediators such as histamine and slow-reacting substance of anaphylaxis (SRS-A). These mediators cause vasodilation and leakage of fluid into the tissues, plus sensory nerve stimulation (leading to itching, sneezing, and cough). Type I is also called immediate-type hypersensitivity because reactions occur within minutes after exposure of a previously sensitized individual to the offending

Table 19.5 Classification of Hypersensitivity Reactions

	Mechanisms		
Type	Induction (Initial Exposure to Antigen)	Elicitation (Re-exposure to Antigen)	Example
I (immediate)	Clonal expansion of B cells; Cytophilic antibody (IgE) generated; binds to mast cells	Antigen binds to cell bound antibody, cross-links receptors, causing release of mediators	Anaphylactic response to bee sting
II (cytolytic)	Clonal expansion of B cells; IgM, IgG generated. Antigen binds to cell membrane	Anamnestic[a] Ig response; binds to cell bound antigen and in the presence of complement or activated macrophages cell lysis occurs	Rh factor incompatibility, Hemolytic anemia in reaction to drug treatment
III (Arthrus)	Clonal expansion of B cells; IgM, IgG generated	Anamnestic Ig response; antigen antibody complexes form in some tissues leading to inflammation	glomerular nephritis, rheumatic heart disease, farmers lung
IV (delayed)	Clonal expansion of antigen-specific T cells occurs	T cells activated, release cytokines, activate macrophages, inflammation	contact dermatitis

[a]Heightened response on re-exposure to antigen.

antigen. Type I reactions include immediate asthmatic responses to allergen, allergic rhinitis (hay fever), atopic dermatitis (eczema), and acute urticaria (hives). The most severe form is systemic anaphylaxis (e.g., in response to a bee sting), which results in anaphylactic shock, and potentially death.

Type II hypersensitivity is the result of antibody-mediated cytotoxicity that occurs when antibodies respond to cell surface antigens. Antibodies bound to antigen on the cell surface activate the complement system and/or macrophages, leading to lysis of the target cell. Frequently blood cells are the targets, as in the case of an incompatible blood transfusion or Rh blood incompatibility between mother and child. The basement membrane of the kidney or lung may also be a target. Autoimmune diseases can result from drug treatments with penicillin, quinidine, quinine, or acetaminophen. Apparently these drugs interact with the cell membrane such that the immune system detects "foreign" antigens on the cell surface. This type of autoimmune disease may also have unknown etiologies.

Type III reactions are the result of antigen-antibody (IgG) complexes that accumulate in tissues or the circulation, activate macrophages and the complement system, and trigger the influx of granulocytes and lymphocytes (inflammation). This is sometimes referred to as the Arthrus reaction and includes postinfection sequelae such as rheumatic heart disease. Farmer's lung, a pneumonitis caused by molds has been attributed to both type III and type IV, and some of the late phase response (4–6 hours after exposure) in asthmatics may be the result of Arthrus-type reactions.

Unlike the preceding three types, type IV, or delayed-type hypersensitivity (DTH), involves T cells and macrophages, not antibodies. Activated T cells release cytokines that cause accumulation and activation of macrophages, which in turn cause local damage. This type of reaction is very important in defense against intracellular infections such as tuberculosis, but is also responsible for contact hypersensitivity responses (allergic contact dermatitis) such as the response to poison ivy. Inhalation of beryllium can result in a range of pathologies, including acute pneumonitis, tracheobronchitis, and chronic beryllium disease, all of which appear to be due to type IV beryllium-specific immune responses. The expression of type IV responses following challenge is delayed, occurring 24 to 48 hours after exposure.

The different types of immune-mediated injury are not mutually exclusive. More than one hypersensitivity mechanism may be involved in the response to a particular antigen. Also the resulting pathology, particularly that caused by type III and IV, reactions may appear very similar, although the mechanisms leading to the effect are different.

19.5 EFFECTS OF CHEMICALS ON ALLERGIC DISEASE

Xenobiotics can affect allergic disease in one of two ways. They can themselves act as antigens and elicit hypersensitivity responses, or they can enhance the development or expression of allergic responses to commonly encountered allergens, such as dust mite. Chemicals that act as allergens include certain proteins that can by themselves induce an immune response and low molecular weight chemicals (known as haptens) that are too small to induce a specific immune response but may react with a protein to induce an immune response that is then hapten specific. Haptens have been associated with both allergic contact dermatitis (ACD), sometimes called contact hypersensitivity

(CHS), and respiratory hypersensitivity. Proteins have been associated with respiratory hypersensitivity and food allergies. When a chemical is an allergen or a hapten, there are two doses of concern, the sensitizing dose and the elicitation dose. In general, the dose required for sensitization is greater than that required to elicit a response in a sensitized individual. Chemicals that enhance the development of allergic sensitization are referred to as adjuvants. Air pollutants have been associated both with enhanced sensitization and exacerbation of allergic respiratory symptoms.

19.5.1 Allergic Contact Dermatitis

Allergic contact dermatitis (ACD) or contact hypersensitivity (CHS) is one of the most common occupational health problems and hence is one of the most common problems associated with immunotoxicity. It is a type IV response that occurs as a result of dermal exposure to chemicals that are haptens. Following dermal exposure, the chemical reacts with host cell protein at the surface of the skin and is picked up by epidermal dendritic cells, known as Langerhans cells. Cytokines released from the epidermal keratinocytes and from Langerhans cells cause maturation and mobilization of the Langerhans cells, which travel to the draining local lymph node and present antigen to lymphocytes. Clonal expansion occurs, enlarging the number of T lymphocytes specific for that allergen and generating memory cells that, in addition to specificity for the allergen, have the propensity to home to the skin. On re-exposure to the chemical, these specific T cells are activated, proliferate, home rapidly to the site of exposure, and produce erythema and edema typical of a type IV response. The reaction to poison ivy is the classic example.

Methods to assess chemicals (drugs, pesticides, dyes, cosmetics, and household products, etc.) for potential to induce CHS are well established, and several protocols using guinea pigs have been in use since the 1950s. These protocols assess the actual disease end point, skin erythema, and edema, following sensitization and challenge with the test agent. Two commonly used tests are the guinea pig maximization test and the Buehler occluded patch test. The sensitization procedure for the maximization tests includes intradermal injection of the test chemical with an adjuvant (intended to enhance the sensitization process) as well as topical application. The Buehler test relies on topical sensitization alone. In both cases, after approximately 2 weeks, animals are challenged at a different site on the skin and erythema and edema are assessed 24 to 48 hours later. This assessment is somewhat subjective and these tests are fairly expensive.

A chemical is considered to be a sensitizer if 30% (maximization) or 15% (Buehler) of the animals respond. Recently a more economical, less subjective, test for CHS has been developed using mice. This test, the local lymph node assay (LLNA), assesses the proliferative response of lymphocytes in the draining lymph node following application of the agent to the ear and is based on our understanding of the immunologic mechanisms underlying CHS; that is, clonal expansion has to occur in the draining lymph node if there is to be allergic sensitization (Figure 19.5). The LLNA is gaining acceptance as a stand-alone alternative to the guinea pig tests and is likely to become the assay of choice.

Finally structure activity approaches have recently been developed to identify contact sensitizers. This approach is based on the concept that the biologic mechanisms that determine a chemical's effect are related to its structure and hence chemicals with similar structures will have similar effects. Computer models have been developed to compare the structure of an unknown chemical to structures in a database for known

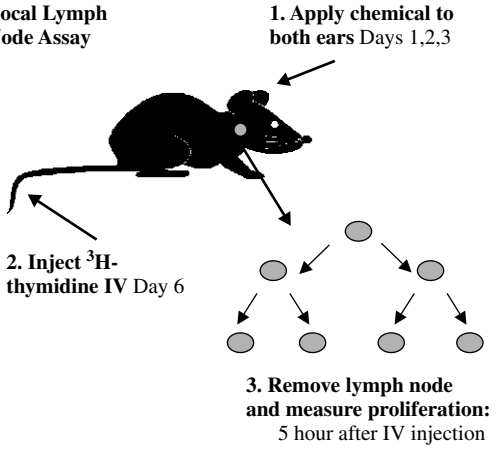

Figure 19.5 Assessing chemicals for potential contact sensitivity. In the local lymph node assay the chemical in question is applied to both ears on three consecutive days. Control mice are treated with vehicle. Radioisotope is injected intravenously on day 6. The draining lymph nodes are removed 5 hours later and the proliferative response is measured by the incorporation of radio isotope. Results are frequently presented as a stimulation index (counts per min (cpm) for the test chemical/cpm for control). (Picture adapted from D. Sailstad, *Lab Animal* **31**: 36, 2002.)

contact sensitizers. CHS lends itself to this approach because there is a large database of chemicals known to cause it, and there is a reasonable understanding of chemical characteristics that facilitate skin penetration, chemical reactivity with host proteins, and immune reactivity.

Because nonspecific inflammatory responses also can occur following chemical exposure to the skin, a distinction must be made between an irritant and a sensitizer. An irritant is an agent that causes local inflammatory effects but induces no immunological memory. Therefore, on subsequent exposures, local inflammation will again result, but there is no enhancement of the magnitude of the response and no change in the dose required to induce the response. In immunologically mediated inflammation (hypersensitivity) there may be no response to a sensitizer during the induction stage, but responses to subsequent exposures are exacerbated. The dose required for elicitation is usually less than that required to achieve sensitization.

19.5.2 Respiratory Allergens

There is evidence that both occupational and environmental exposures to chemicals (both proteins and haptens) can result in the induction or exacerbation of respiratory allergies (Table 19.6). Of particular concern is the induction of allergic asthma. In sensitized asthmatic individuals the antigen challenge generally causes a type I (IgE-mediated) immediate hypersensitivity response with release of mediators responsible for bronchoconstriction. Between 2 and 8 hours after the immediate response, asthmatics experience a more severe and prolonged (late phase) reaction that is characterized by mucus hypersecretion, bronchoconstriction, airway hyperresponsiveness to a variety of nonspecific stimuli (e.g., histamine, methacholine), and airway inflammation characterized by eosinophils. This later response is not mediated by IgE.

Table 19.6 Example of Chemicals Associated with Respiratory Allergy

Proteins
Enzymes
Latex
Animal dander
Dust mite
Molds
Cockroach
Microbial pesticides
Low molecular weight (<3000)-haptens
Toluene diisocyanate
Diphenylmethane diisocyanate
Phthalic anhydride
Trimellitic anhydride
Platinum salts
Reactive dyes
Adjuvants
Ozone
Nitrogen dioxide
Diesel exhaust
Residual oil fly ash

Although proteins are generally immunogens, not all proteins are allergens and there is a range of potencies for those that are. There is also a strong genetic component associated with susceptibility to develop allergic reactions to proteins. Susceptible individuals are called atopic. There is at present no structural motif that can be used to characterize a protein as an allergen for hazard identification. Examples of occupational protein exposures associated with respiratory allergy and asthma include enzymes, latex, flour (both the grain itself and fungal contaminants), and animal dander. Environmental (mostly indoor) exposure including molds, spores, dust mite, animal dander, and cockroach have also been associated with this type of respiratory disease. Because this is a type 1 response, cytophilic antibodies (IgE) specific for the allergen are frequently used to identify proteins that may cause this effect. For example, in order to determine the etiology of occupational asthma in human subjects, the skin prick test is often used. Different proteins are injected under the skin to test for the presence of cytophilic antibodies in order to identify which proteins are causing a response in an individual. Serum may also be tested for protein specific IgE. Because IgE can sometimes be detected in the absence of respiratory responses, a positive IgE test may be followed by an assessment of respiratory responses.

Under very controlled situations patients may be exposed via the respiratory route to suspect allergens (broncho- provocation test) and respiratory function monitored to pinpoint the offending allergen. Guinea pigs and mice have been used to test proteins for potential allergenicity. Animals are usually sensitized by the respiratory route and monitored for the development of cytophilic antibody (IgG1 in guinea pigs, IgE in mice) as well as increased respiratory rate and other changes in pulmonary function.

The guinea pig intratracheal test has been used to establish the relative potency of different detergent enzymes and establish safe occupational exposure levels. As the name implies, guinea pigs are sensitized by intratracheal exposure and induction of cytophilic antibodies are assessed. Dose responses obtained for new enzymes are compared to a reference enzyme for which safe exposure levels have been established. The relative potency of the new enzyme to this reference is used to establish a safe exposure level for the new enzyme.

Exposure to certain low (<3000) molecular weight compounds (haptens) has also been associated with the development of occupational asthma. Highly reactive compounds such as the diisocyanates or acid anhydrides have the capacity to react with protein and induce an immune response. Toluene diisocyanate (TDI) and trimellitic anhydride are the compounds that have been most extensively studied in this regard. There is a great deal of interest in developing a test to screen chemicals for this type of effect in order to avoid induction of immune responses that could lead to occupational asthma. Although specific IgE antibodies have been detected in some individuals with TDI asthma, it has not been uniformly present and some of these individuals exhibit the late phase but not the immediate response. Hence, unlike proteins, there is less certainty about the mechanisms underlying respiratory allergic responses to low molecular weight compounds.

Structure activity approaches similar to those described for contact sensitizers have been developed, but this approach has limitations because the database of known respiratory sensitizers is small compared to contact sensitizers and the underlying mechanisms are less well defined. At the other extreme guinea pigs have been exposed by inhalation for a number of days, rested, and then challenged at a later date by inhalation with subsequent monitoring of respiratory responses. Although this approach has produced a good model of TDI asthma, it is too cumbersome and expensive for routine testing. Because the capacity to interact with protein is a pre-requisite to allergenicity, it has been suggested that testing for protein reactivity in vitro could provide an initial screening test for chemicals. Also, because it appears that respiratory sensitizers are a subset of chemicals that produce positive results in a contact sensitivity test, it has been suggested that the LLNA test be used as the first tier in screening chemicals for this effect. The problem then becomes separating chemicals that are strictly contact sensitizers from those that have the capacity to cause respiratory sensitization. Efforts have been made to determine whether differences in responses to dermal application of these chemicals could provide a means for making this distinction. One proposal is to assess total serum IgE following dermal exposure, assuming that respiratory sensitizers would produce a bigger IgE signal. Another approach has been to assess cytokine profiles in the draining lymph node following dermal exposure. Different subsets of T helper (Th) cells, have been associated with type I immediate (Th2) and type IV delayed (Th1) responses. These different populations of T cells are distinguished by different cytokine profile and efforts are underway to use these differing profiles to distinguish respiratory from contact sensitizers. However, there is as yet no well-validated, well-accepted test to assess low molecular weight chemicals for the capacity to induce respiratory allergy. This remains a subject of intense research.

19.5.3 Adjuvants

An adjuvant is a compound administered in conjunction with an antigen that nonspecifically enhances the immune response to that antigen. Adjuvants are used in

vaccines to promote immunogenicity. There is now growing concern that chemicals in our environment (particularly, air pollutants) might act as adjuvants for allergic sensitization to common allergens such as dust mite and pollen. Laboratory rodents have been used to show that nitrogen dioxide, residual oil fly ash, and diesel exhaust enhance allergic sensitization and disease. Enhanced sensitization to an allergen has also been demonstrated in rhesus monkeys exposed to ozone and humans exposed to diesel exhaust. The significance of these findings in terms of enhanced burden of respiratory allergies in the human population is unclear. As in other areas of toxicology, simultaneous environmental exposures to agents that are not the agent of immediate concern can certainly influence outcomes. Adjuvancy is a concern that likely extends beyond air pollution and type 1 responses.

19.6 EMERGING ISSUES: FOOD ALLERGIES, AUTOIMMUNITY, AND THE DEVELOPING IMMUNE SYSTEM

There are several emerging issues in immunotoxicology. These active areas of research will be only briefly described here because there are currently more questions than answers.

Toxicologists have recently been drawn into the area of food allergy by advances in biotechnology and the need to assess the safety of genetically modified foods in terms of potential allergenicity. There is concern that insertion of a novel gene into a food crop (e.g., to increase yield or pest resistance) might inadvertently introduce a new allergen into the food supply. Food allergies are relatively rare, affecting approximately 5% of children and 2–3% of adults, and even in these individuals, most proteins are not food allergens. However, when food allergy does occur, the consequences can be severe. Anaphylactic (life-threatening) reactions to peanuts provide the best example. Unfortunately, the mechanisms underlying food allergies (or the mechanisms that protect most of people from developing reactions to the foreign proteins they eat), the characteristics that make a protein a food allergen, and the characteristics that make an individual susceptible to food allergies are poorly understood at this time. These are some of the issues that need to be resolved in order to develop appropriate safety assessment tools.

Autoimmune diseases affect about 3% of the population and comprise a diverse array of both organ specific (e.g., type I diabetes, thyroiditis) and systemic (systemic lupus erythematosis) diseases. Susceptibility includes a strong genetic component, and in some cases women appear to be more vulnerable than men. Xenobiotics might affect the development or progression of autoimmune disease. A variety of mechanisms could contribute to xenobiotic effects on the development and maintenance of immune tolerance or unmasking or modification of self proteins. There is evidence that exposure to certain drugs, heavy metals, silica, and endocrine disruptors are a concern in this regard. Current research includes both human and animals studies to determine the extent of risk and ways to assess and control it.

Finally there is growing concern that the developing immune system may be particularly vulnerable to xenobiotic exposures and that perinatal and/or in utero exposures may have a lifelong impact on susceptibility to infectious, allergic, or autoimmune disease. As in other areas of toxicology, tests designed to assess the risk of immunotoxicity for adults may not be sufficient to protect children and research is currently underway to determine how best to meet this need.

Clearly, exposure to xenobiotics can have a number of effects on the immune system that in turn can affect an array of health outcomes. In some areas of immunotoxicology significant progress has been made in terms of identifying and understanding the risks associated with xenobiotic exposure. In other areas more research is needed.

SUGGESTED READING

Benjamini, E., G. Sunshine, and S. Leskowitz. *Immunology A Short Course*, 4th ed. New York: Wiley, 2000.

Gilmour, M. I., M. J. K. Selgrade, and A. L. Lambert. Enhanced allergic sensitization in animals exposed to particulate air pollution. *Inhalation Toxicol.* **12** (suppl. 3): 373–380, 2000.

Kimber, I., I. L. Bernstein, M. H. Karol, M. K. Robinson, K. Sarlo, and M. J. K. Selgrade. Workshop overview: Identification of respiratory allergens. *Fundam. Appl. Toxicol.* **33**: 1–10, 1996.

Luster, M. I., A. E. Munson, P. T. Thomas, M. P. Holsapple, J. D. Fenters, K. L. White, L. E. Lauer, D. R. Germolec, G. J. Rosenthal, and J. H. Dean. Development of a testing battery to assess chemical-induced immunotoxicity: National Toxicology Program's guidelines for immunotoxicity evaluation in mice. *Fundam. Appl. Toxicol.* **10**: 2–19, 1988.

Luster, M. I., et al. Risk assessment in immunotoxicology I. Sensitivity and predictability of immune tests. *Fundam. Appl. Toxicol.* **18**: 200–210, 1992, and II. Relationships between immune and host resistance tests. *Fundam. Appl. Toxicol.* **21**: 71–82, 1993.

Metcalfe, D. E., et al. Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Crit. Rev. Food Sci. Nutr.* **36S**: S165–S186, 1996.

Roitt, I. *Roitt's Essentials in Immunology*. London: Blackwell, 2001.

Sailstad, D. M. Murine local lymph node assay: An alternative test method for skin hypersensitivity testing. *Lab. Animal* **31**: 36–41, 2002.

Sarlo, K. Human Health Risk Assessment: Focus on Enzymes. In *Proceeding of the 3rd World Conference on Detergents*, A. Cahn, ed. Champaign, I: Am. Oil Chem. Soc. Press, 1994, pp. 54–57.

CHAPTER 20

Reproductive System

STACY BRANCH

20.1 INTRODUCTION

What is reproductive toxicity? Reproductive toxicity refers to any adverse effect on any aspect of male or female sexual structure, function, and lactation including effects on the reproductive potential and viability of the offspring. This concept may also include the following:

Organ toxicity. Can interfere with normal system function.

Teratogenicity. Ability to cause dysmorphogenesis in the developing fetus.

Behavioral teratogenicity. Ability to adversely affect the mental development of the fetus.

Developmental toxicity. Includes both teratogenicity terms above and abnormal postnatal development.

This chapter discusses both male and female reproductive toxicity and provides an overview of reproductive physiology.

20.2 MALE REPRODUCTIVE PHYSIOLOGY

The anterior pituitary is stimulated by the hypothalamus (via the gonadotrophic hormone releasing hormone) to release gonadotrophic hormones (leutenizing hormone—LH; and follicle stimulating hormone—FSH). In addition to LH and FSH, prolactin is released by the anterior pituitary. The target of LH and FSH in the male is the testis. While LH stimulates steroidogenesis, FSH has its primary effects on the sertoli cells. The role of prolactin (which is inhibited by dopamine) is to modulate the effects of LH in the testicular tissue. Critical points within the hypothalamic–pituitary–gonadal axis (Figure 20.1) may be susceptible to alterations by xenobiotics, leading to altered reproductive function and pathology.

A Textbook of Modern Toxicology, Third Edition, edited by Ernest Hodgson
ISBN 0-471-26508-X

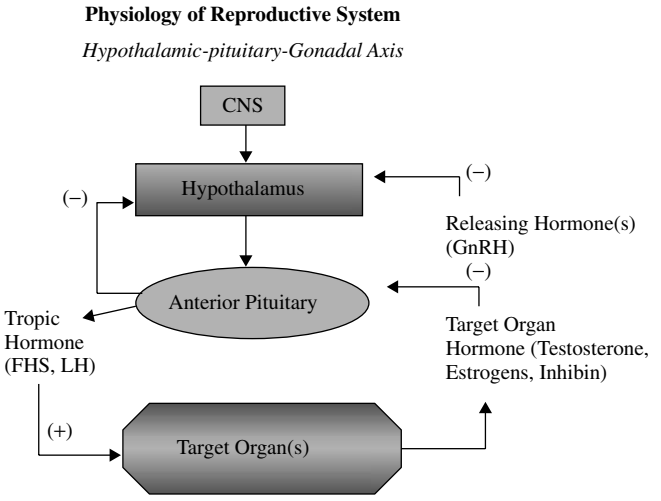

Figure 20.1 Hypothalamic–pituitary–gonadal axis. Negative feedback is designated by (−), and positive feedback is designated by (+).

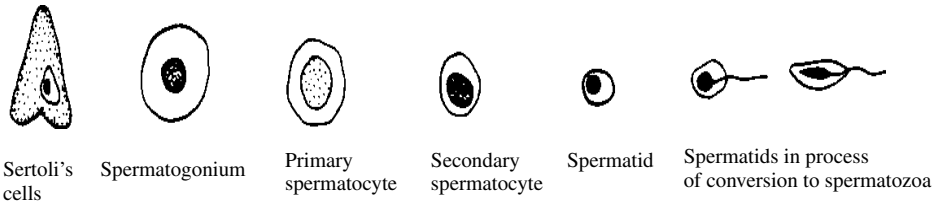

Figure 20.2 Spermatogenesis. (Adapted from J. A. Thomas, in *Casarett and Doull's Toxicology: The Basic Science of Poisons*, 6th ed., C. D. Klassen, ed., McGraw-Hill, 2001.)

Two major components of the testis are the seminiferous tubules (site of spermatogenesis) and the interstitial compartment. The interstitial compartment contains Leydig cells, which produce testosterone under the influence of LH. Androgens control spermatogenesis (Figure 20.2), growth and activity of accessory sex glands, masculinization, male behavior, and various metabolic functions. Secretion of androgen by the developing fetal testis is essential for differentiation of the gonads, which includes regression of Müllerian ducts and the development of Wolffian ducts.

20.3 MECHANISMS AND TARGETS OF MALE REPRODUCTIVE TOXICANTS

20.3.1 General Mechanisms

Toxicants may mimic endogenous compounds (i.e., hormones), thus acting as agonists or antagonists. Toxicants may be directly cytotoxic or may be activated to toxic compounds. Some toxicants may have indirect effects by inhibiting key enzymes involved

in steroid synthesis. Below are examples of how selected toxicants affect various susceptible targets of the male reproductive tract.

20.3.2 Effects on Germ Cells

Epidemiological data have indicated that wives of men exposed to the compound vinyl chloride experience spontaneous abortion. Vinyl chloride is used to manufacture polyvinyl chloride (PVC). PVC is a component of various plastic products, including pipes, furniture, and automobile upholstery. Ethylnitrosourea (ENU) is a mutagenic agent that has been used extensively in mouse mutagenesis. ENU acts on male spermatogonial stem cells, introducing mutations. It is also capable of inducing reversible sterility in mice. Actinomycin D is an older chemotherapeutic drug that has been used in cancer therapy for many years. Actinomycin D is commonly used in the treatment of gestational trophoblastic cancers, testicular cancer, Wilm's tumor, and rhabdomyosarcoma. Actinomycin D intercalates with DNA and disrupts the structure and function of the DNA of the spermatozoa.

20.3.3 Effects on Spermatogenesis and Sperm Quality

The antibiotics and antimetabolites used in cancer treatment (i.e., vinblastine) are spermatotoxic and affect semen quality. Ionizing radiation is also spermatotoxic (with spermatogonial cells being the most sensitive). Prolonged scrotal heating is a factor that affects the earlier states of spermatogenesis.

20.3.4 Effects on Sexual Behavior

Anabolic steroids, antidepressants and drugs of abuse affect libido, potency, and ejaculatory function. Anabolic steroids are derivatives of testosterone, and have strong genitotropic effects. There is published evidence indicating that anabolic steroids increases sexual desire; however, the frequency of erectile dysfunction is also increased. Treatment with the antidepressant fluoxetine has been associated with sexual side effects including delayed or nonexistent ejaculation and hyposexuality. Mice treated in utero with the anitleukemic agent 5-aza-2′-deoxycytidine exhibit abnormal reproductive behavior and low reproductive capacity.

20.3.5 Effects on Endocrine Function

Cimetidine (for treatment of peptic ulcers) competes with dihydrotestosterone for receptors in the testis and accessory sex glands. The more common sequelae are low sperm count and gynacomastia. Epidemiological evidence has shown that occupation exposure to oral contraceptives can induce gynacomastia in exposed males. Diethylstilbestrol (DES) antagonizes the activity of fetal testosterone. In the male offspring, testicular hypoplasia, abnormal semen parameters, and infertility result. Ketoconazole has be shown to be transported to the seminal fluid and to immobilize the sperm.

20.4 FEMALE REPRODUCTIVE PHYSIOLOGY

As described previously for the male, the female hormonal signaling is composed of four primary levels: CNS, hypothalamus, anterior pituitary, and gonads. The gonadotropin-releasing hormones of the hypothalamus stimulates the anterior pituitary to release LH and FSH. Subsequently LH and FSHS stimulate the release of estrogen and progesterone from the ovaries (Figure 20.3). Estrogen is secreted in the growing follicle and has effects on the uterus. The oocytes are formed before birth, then develop into the primary oocytes after meiosis. At the time of puberty, the release of gonadotropin stimulates the oocytes to develop into graafian follicles (Figure 20.3).

Ovarian Cycles. Estrus is the period when the female mammal is most receptive to the male (coincides with high levels of circulating estrogen). Rodents are considered to be polyestrous and have a succession of estrus cycles. Cats are seasonally (spring, early fall) polyestrus, while dogs are monestrous. Humans and higher primates cycle at monthly intervals. Although most mammals ovulate spontaneously, some mammals (cats and minks) undergo provoked or induced ovulation (i.e., stimulated by mating). The estrus cycle and the resulting differences in circulating hormone concentrations at different stages of the cycle are depicted in Figure 20.4. The changes in circulating hormone and the stage of follicle development during an adverse toxicant insult results in a variety of toxicological manifestations.

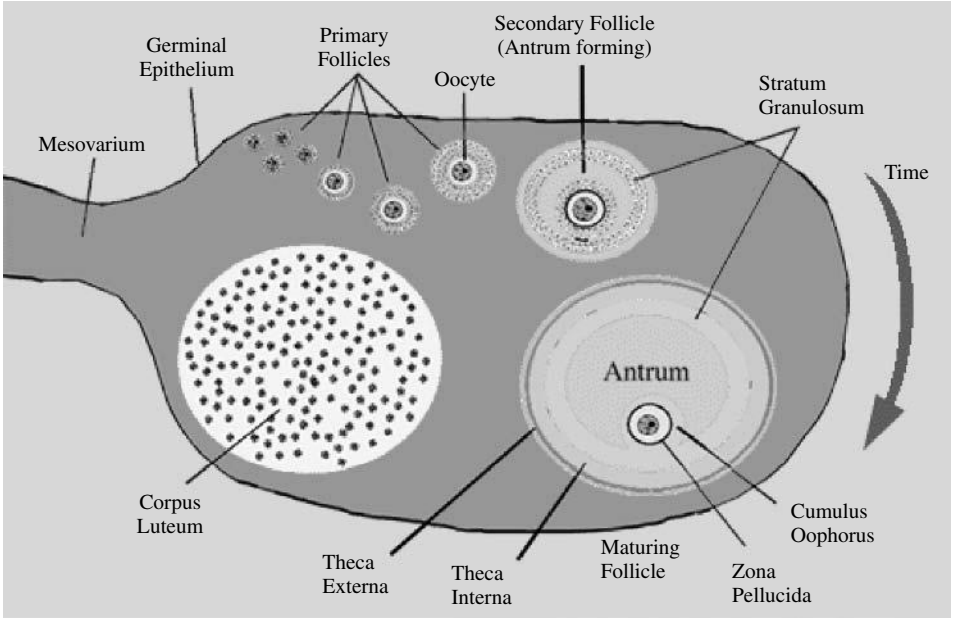

Figure 20.3 The arrow follows the ovarian follicles (time course) from their maturation from primary follicles to the corpus luteum. (Adapted from Web site of Dr. Steven Scadding and Dr. Sandra K. Ackerley, http://www.uoguelph.ca/zoology/devobio/210labs/ovary4.html.)

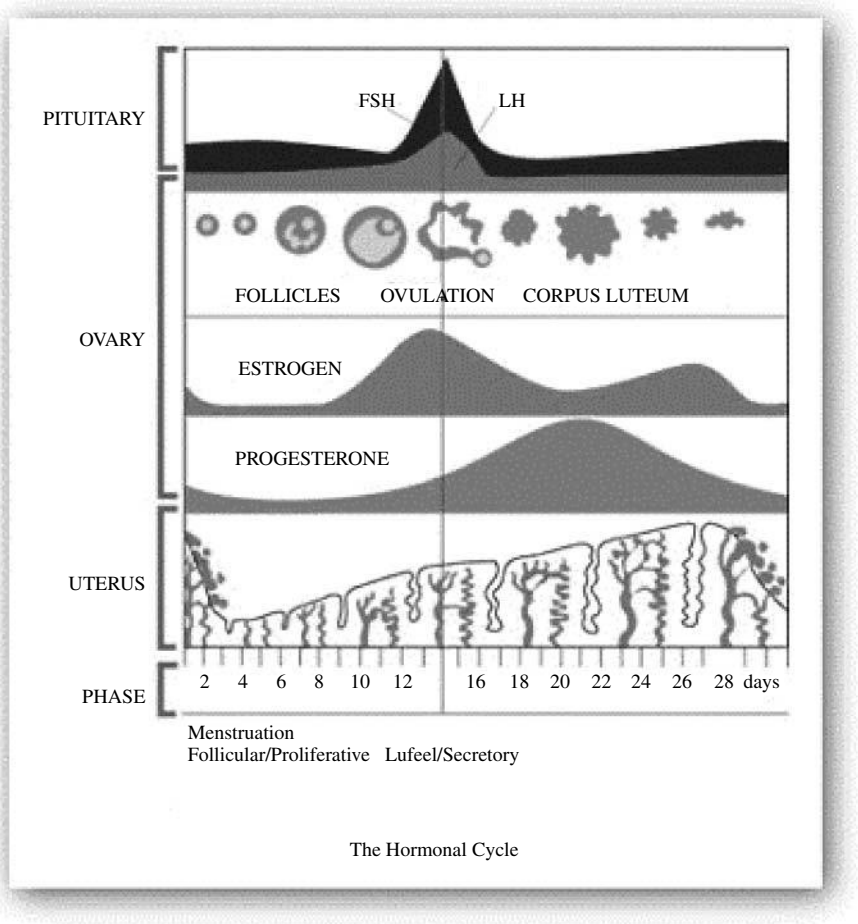

Figure 20.4 Women's menstrual cycle. (Courtesy of Women's Health Interactive, http://www.womens-health.com/health_center/gynecology/gyn_repro_menstrual.html.)

20.5 MECHANISMS AND TARGETS OF FEMALE REPRODUCTIVE TOXICANTS

Xenobiotics can adversely affect the normal functions of the cells/organs of the reproductive system. These agents may induce a variety of outcomes, including prevention of ovulation and impairment of ovum transport, fertilization, or implantation. Endocrine disruptors may mimic endogenous hormones as well as directly destroy cellular components, leading to cell death. More indirect effects may include inhibition of key enzymes involved in steroid synthesis.

20.5.1 Tranquilizers, Narcotics, and Social Drugs

Compounds within this class of substances can inhibit hypothalamic–pituitary–ovarian axis function by inhibiting gonadotropin secretion. Subsequently ovulation and estrus is suppress leading to infertility or reduced fertility.

20.5.2 Endocrine Disruptors (EDs)

Endocrine disruptors are compounds (synthetic and naturally occurring) that can alter the normal hormonal balance and function in animals. The historical ED diethylstilbestrol (DES) is a classic example of an endocrine disrupter affecting female reproductive health. In utero exposure of females to DES is associated with the induction of vaginal carcinomas apparent after puberty. In experimental mice, estrogenic substances cause accelerated sexual maturation and irregular estrous cycles and prolonged estrous. In rats, xenoestrogens such as kepone and methoxychlor cause masculinization of the exposed female rats. These rats do not ovulate, lack stimulation of the LH surge, and exhibit male sexual behavior. In humans, estrogen mimicking compounds can alter natural hormonal cycles and have been associated with breast cancer induction. Certain environmental EDs may function as promoters or inducers of carcinogenesis. Polychlorinated biphenyls (PCBs) and a trichloroethane compound (DDT) are persistent in the environment. Serum DDE (a DDT metabolite) levels have been found to correlate with breast cancer incidence.

20.5.3 Effects on Germ Cells

As previously described for the male reproductive toxicity, the class of toxicants affecting germ cells can alter the structure of genetic material (chromosomal aberrations, alterations in meiosis, DNA synthesis, and replication). Mature oocytes have a DNA repair capacity different from that of mature sperm, but this capacity decreases at the period of meiotic maturation.

20.5.4 Effects on the Ovaries and Uterus

Cyclophosphamide and vincristine are examples of alkylating agents capable of inducing gonadal dysfunction. Premature menopause is a primary outcome of exposure to these agents. Amenorrhea and abnormal hormonal levels are characteristics of the ovarian dysfunction induced by cyclophosphamide.

Premature ovarian failure can be induced in offspring exposed in utero by active metabolites such as 6-mercaptopurine. Tamoxifen (treatments for breast cancer) and clomiphene (to induce ovulation) are antiestrogens that can inhibit uterine decidual induction in pseudopregnant rats.

20.5.5 Effects on Sexual Behavior

Normal sexual activity is associated with ovulation in most female mammals. Compounds affecting this process can adversely affect female libido. Ovarian failure induced by xenobiotic compounds has been associated with a decrease in libido in women. Certain types of oral contraceptives as well as drugs of abuse (methadone, cannabis, alcohol) cause decreases in female libido. The treatment for hirsutism, excessive growth of hair in both normal and abnormal locations, is the compound cyproterone acetate. It is an antiandrogen that has the side effect of severely decreasing libido in women.

SUGGESTED READING

Klaassen, C. D. *Cassarett and Doull's Toxicology: The Basic Science of Poisons*, 6th ed. New York: McGraw-Hill, 2001.

Korach, K. S. *Reproductive and Developmental Toxicology*. New York: Dekker, 1998.

Naz, R. K. *Endocrine Disruptors*. Boca Raton, FL: CRC Press, 1997.

Ballantyne, B., T. Marrs, and P. Turner, eds. *General and Applied Toxicology*, college ed., New York: Macmillan, 1995.

PART VI

APPLIED TOXICOLOGY

CHAPTER 21

Toxicity Testing

HELEN CUNNY and ERNEST HODGSON

21.1 INTRODUCTION

Although testing for toxicity, usually for the purposes of human health risk assessment, might be expected to be one of the more routine aspects of toxicology, it is actually one of the more controversial. Among the many areas of controversy are the use of animals for testing and the welfare of the animals, extrapolation of animal data to humans, extrapolation from high-dose to low-dose effects, and the increasing cost and complexity of testing protocols relative to the benefits expected. New tests are constantly being devised and are often added to testing requirements already in existence.

Most testing can be subdivided into in vivo tests for acute, subchronic, or chronic effects and in vitro tests for genotoxicity or cell transformation, although other tests are used and are described in this chapter. Any chemical that has been introduced into commerce or that is being developed for possible introduction into commerce is subject to toxicity testing to satisfy the regulations of one or more regulatory agencies. Furthermore compounds produced as waste products of industrial processes (e.g., combustion products) are also subject to testing.

Toxicity assessment is the determination of the potential of any substance to act as a poison, the conditions under which this potential will be realized, and the characterization of its action. *Risk assessment*, however, is a quantitative assessment of the probability of deleterious effects under given exposure conditions. Both are involved in the regulation of toxic chemicals. *Regulation* is the control, by statute, of the manufacture, transportation, sale, or disposal of chemicals deemed to be toxic after testing procedures or according to criteria laid down in applicable laws.

Testing in the United States is carried out by many groups: industrial, governmental, academic, and others. Regulation, however, is carried out by a narrow range of governmental agencies, each charged with the formulation of regulations under a particular law or laws and with the administration of those regulations. The principal regulatory agencies for the United States are shown in Table 21.1. Other industrialized countries have counterpart laws and agencies for the regulation of toxic chemicals.

A Textbook of Modern Toxicology, Third Edition, edited by Ernest Hodgson
ISBN 0-471-26508-X

Table 21.1 Some Agencies and Statutes Involved in Regulation of Toxic Chemicals in the United States

Food and Drug Administration (FDA)
Food, Drug, and Cosmetic Act
Labor Department
Occupational Safety and Health Act
Consumer Product Safety Commission
Consumer Product Safety Act
Environmental Protection Agency (EPA)
Federal Insecticide, Fungicide, and Rodenticide Act
Clean Air Act
Federal Water Pollution Control Act
Safe Drinking Water Act
Toxic Substances Control Act
Resource Conservation and Recovery Act
State governments
Various state and local laws
Enforcement of certain aspects of federal law delegated to states

Although the objective of much, but by no means all, toxicity testing is the elimination of potential risks to humans, most of the testing is carried out on experimental animals. This is necessary because our current knowledge of quantitative structure activity relationships (QSAR) does not permit accurate extrapolation to new compounds. Human data are difficult to obtain experimentally for ethical reasons but necessary for such deleterious effects as irritation, nausea, allergies, odor evaluation, and some higher nervous system functions. Some insight may be obtained in certain cases from occupational exposure data, although this often tends to be irregular in time and not clearly defined as to the composition of the toxicant or the exposure levels, because multiple exposure is common. Clearly, any experiments involving humans must be carried out under carefully defined conditions after other testing is complete.

Although for a variety of reasons extrapolation from experimental animals to humans presents problems, including differences in metabolic pathways, dermal penetration, mode of action, and others, experimental animals present numerous advantages in testing procedures. These advantages include the possibility of clearly defined genetic constitution and their amenity to controlled exposure, controlled duration of exposure, and the possibility of detailed examination of all tissues following necropsy.

Although not all tests are required for all potentially toxic chemicals, any of the tests shown in Table 21.2 may be required by the regulations imposed under a particular law. The particular set of tests required depends on the predicted or actual use of the chemical, the predicted or actual route of exposure, and the chemical and physical properties of the chemical.

Table 21.2 Summary of Toxicity Tests and Related End Points

I Chemical and physical properties
For the chemical in question, probable contaminants from synthesis as well as intermediates and waste products from the synthetic process

II Exposure and environmental fate
A. Degradation studies—hydrolysis, photodegradation, etc.
B. Degradation in soil, water, under various conditions
C. Mobility and dissipation in soil, water, and air
D. Accumulation in plants, aquatic animals, wild terrestrial animals, food plants, and animals, etc.

III In vivo tests
A. Acute
1. LD50 and LC50—oral, dermal or inhaled
2. Eye irritation
3. Dermal irritation
4. Dermal sensitization
B. Subchronic
1. 30- to 90-day feeding
2. 30- to 90-day dermal or inhalation exposure
C. Chronic/reproduction
1. Chronic feeding (including oncogenicity tests)
2. Teratogenicity
3. Reproduction (multi-generation)
D. Special tests
1. Neurotoxicity
2. Potentiation
3. Metabolism
4. Pharmacodynamics
5. Behavior

IV In vitro tests
A. Mutagenicity—prokaryote (Ames test)
B. Mutagenicity—eukaryote (*Drosophila*, mouse, etc.)
C. Chromosome aberration (*Drosophila*, sister chromatid exchange, etc.)

V Effects on wildlife
Selected species of wild mammals, birds, fish, and invertebrates: acute toxicity, accumulation, and reproduction in laboratory-simulated field conditions

21.2 EXPERIMENTAL ADMINISTRATION OF TOXICANTS

21.2.1 Introduction

Regardless of the chemical tested and whether the test is for acute or chronic toxicity, all in vivo testing requires the reproducible administration of a known dose of the chemical under test, applied in a reproducible manner, that is generally related to the expected route of humans exposure. The nature and degree of the toxic effect can be

Table 21.3 Variation in Toxicity by Route of Exposure

Chemical	Species/Gender	Route	LD50 (mg/kg)
N-Methyl-*N*-(1-naphthyl) fluoroacetamide[a]	Mouse/M	Oral	371
		Dermal	402
		Subcutaneous	250
	Rat/M	Oral	115
		Dermal	300
		Subcutaneous	78
Chlordane[b]	Rat/M	Oral	335
		Dermal	840
Endrin[b]	Rat/M	Oral	18
		Dermal	18

[a] Data from Y. Hashimoto, et al., *Tox. Appl. Pharmacol.* **12**: 536–547, 1968.
[b] Data from J. R. Allen, et al., *Pharmacol. Therap.* **7**: 513–547, 1979.

affected by the route of administration (Table 21.3). This may be related to differences at the portals of entry or to effects on pharmacokinetic processes. In the latter case, one route (e.g., intravenous) may give rise to a concentration high enough to saturate some rate-limiting process, whereas another (e.g., subcutaneous) may distribute the dose over a longer time and avoid such saturation. Another key question is that of appropriate controls. To identify effects of handling and other stresses as well as the effects of the solvents or other carriers, it is usually better to compare treated animals with both solvent-treated and untreated or possibly sham-treated controls.

21.2.2 Routes of Administration

Oral. Oral administration is often referred to as administration per os (PO). Compounds can be administered mixed in the diet, dissolved in drinking water, by gastric gavage, by controlled-release capsules, or by gelatin capsules. In the first two cases, either a measured amount can be provided or access can be ad libitum (available 24 hours per day), with the dose estimated from consumption measurements. For certain tests pair-feeding of controls should be considered; that is, controls are permitted only the amount of food consumed by treated animals. In any case it is essential to consider possible nutritional effects caused by reduction of food intake due to distasteful or repellent test materials. In gastric gavage, the test material is administered through a stomach tube or gavage needle; if a solvent is necessary for preparation of dosing solutions or suspensions, the vehicle is administered also to control animals.

Dermal. Dermal administration is required for estimation of toxicity of chemicals that may be absorbed through the skin, as well as for estimation of skin irritation and photosensitization. Compounds are applied, either directly or in a suitable solvent, to the skin of experimental animals after hair has been removed by clipping. Often dry materials are mixed with water to make a thick paste that can be applied in a manner that ensures adequate contact with the skin. Frequently the animals must be restrained to prevent licking and hence oral uptake of the material. Solvent and restraint controls should be considered when stress is involved. Skin irritancy tests may be conducted on either animals or humans, using volunteer test panels for human tests.

Inhalation. The respiratory system is an important portal of entry, and for evaluation purposes animals must be exposed to atmospheres containing potential toxicants. The generation and control of the physical characteristics of such contaminated atmospheres is technically complex and expensive in practice. The alternative—direct instillation into the lung through the trachea—presents problems of reproducibility as well as stress and for these reasons is generally unsatisfactory.

Inhalation toxicity studies are conducted in inhalation chambers. The complete system contains an apparatus for the generation of aerosol particles, dusts, or gas mixtures of defined composition and particle size, a chamber for the exposure of experimental animals, and a sampling apparatus for the determination of the actual concentration within the chamber. All these devices present technical problems that are difficult to resolve. For rat studies, a particle size of 4 microns is usually targeted.

Animals are normally exposed for a fixed number of hours each day and a fixed number of days each week. Exposure may be nose only, in which the nose of the animal is inserted into the chamber through an airtight ring; or whole body, in which animals are placed inside the chamber. In the latter case, variations due to unequal distribution in the test atmosphere are minimized by rotation of the position of the cages in the chamber during subsequent exposures. Whole body exposure results are usually less satisfactory due to test material accumulation on the fur of the animals and subsequent ingestion during grooming. Figure 21.1 shows a typical inhalation system and supporting equipment.

Injection. Except for certain pharmaceuticals and drugs of abuse, injection (parenteral administration) does not correspond to any of the expected modes of exposure. Injection may be useful, however, in mechanistic studies or in QSAR studies in order to bypass absorption and/or permit rapid action. Methods of injection include intravenous (IV), intramuscular (IM), intraperitoneal (IP), and subcutaneous (SC). Infusion of test

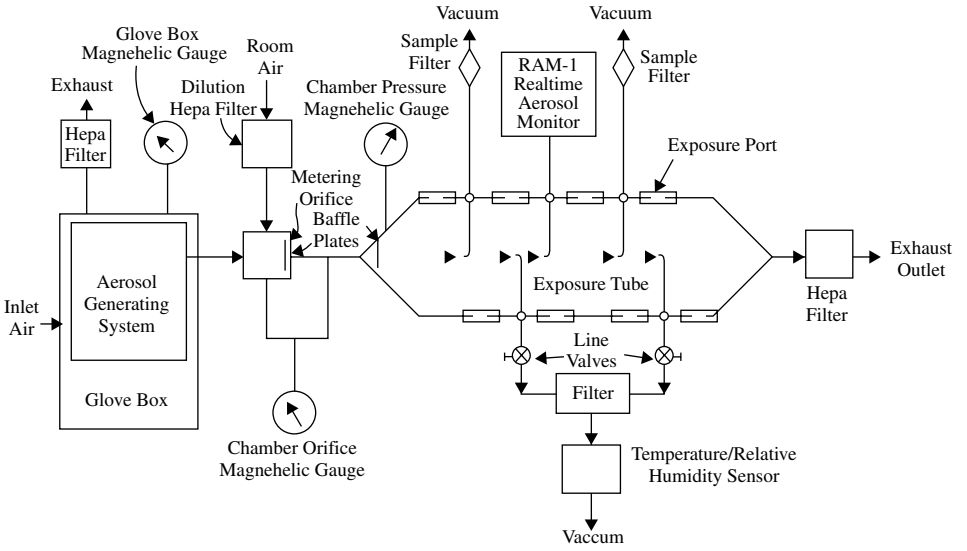

Figure 21.1 An inhalation exposure system (Modified from Adkins, B. et al. *Am. Int. Hyg. Assoc. J.* **41**: 494, 1980.)

materials over an extended period is also possible. Again, both solvent controls and untreated controls are necessary for proper interpretation of the results.

21.3 CHEMICAL AND PHYSICAL PROPERTIES

Although the determination of chemical and physical properties of known or potential toxicants does not constitute a test for toxicity, it is an essential preliminary for such tests.

The information obtained can be used in many ways. Structure activity comparisons with other known toxicants may indicate the most probable hazards. These comparisons may also aid identification in subsequent poisoning episodes. Determination of stability to light, heat, freezing, and oxidizing or reducing agents may enable preliminary estimates of persistence in the environment as well as indicate the most likely breakdown products that may also require testing for toxicity. Establishing such properties as the lipid solubility or octanol/water partition coefficient may enable preliminary estimates of rate of uptake and persistence in living organisms. Vapor pressure may indicate whether the respiratory system is a probable route of entry. Knowledge of the chemical and physical properties must be acquired in order to develop analytical methods for the measurement of the compound and its degradation products. If the chemical is being produced for commercial use, similar information is needed on intermediates in the synthesis or by-products of the process because both are possible contaminants in the final product.

21.4 EXPOSURE AND ENVIRONMENTAL FATE

Data on exposure and environmental fate are needed, not to determine toxicity, but to provide information that may be useful in the prediction of possible exposure in the event that the chemical is toxic. These tests are primarily useful for chemicals released into the environment such as pesticides, and they include the rate of breakdown under aerobic and anaerobic conditions in soils of various types, the rates of leaching into surface water from soils of various types, and the rate of movement toward groundwater. The effects of physical factors on degradation through photolysis and hydrolysis studies and the identification of the product formed can indicate the rate of loss of the hazardous chemical or the possible formation of hazardous degradation products. Tests for accumulation in plants and animals and movement within the ecosystem are considered in Section 21.7.

21.5 IN VIVO TESTS

Traditionally the basis for the determination of toxicity has been administration of the test compound, in vivo, to one or more species of experimental animal, followed by examination for clinical signs of toxicity and/or mortality in acute tests. In addition pathological examination for tissue abnormalities is also performed, especially in tests of longer duration. The results of these tests are then used by a variety of extrapolation

techniques to estimate hazard to humans. These tests are summarized in the remainder of this section. While these tests offer many advantages and are widely used, they suffer from a number of disadvantages: they require the use of animals, whose numbers are often deemed unnecessary by both animal rights and animal welfare advocates; they are extremely expensive to conduct, and they are time-consuming. As a result they have been supplemented by many specialized in vitro tests, some of which are summarized in Section 21.6, and research is ongoing to further develop tests with fewer disadvantages.

21.5.1 Acute and Subchronic Toxicity Tests

Acute toxicity test methods measure the adverse effects that occur within a short time after administration of a single dose of a test substance. This testing is performed principally in rodents and is usually done early in the development of a new chemical or product to provide information on its potential toxicity. This information is used to protect individuals who are working with the new material and to develop safe handling procedures for transport and disposal. The information gained also serves as the basis for hazard classification and labeling of chemicals in commerce. Acute toxicity data can help identify the mode of toxic action of a substance and may provide information on doses associated with target-organ toxicity and lethality that can be used in setting dose levels for repeated-dose studies. This information may also be extrapolated for use in the diagnosis and treatment of toxic reactions in humans. The results from acute toxicity tests can provide information for comparison of toxicity and dose-response among members of chemical classes and help in the selection of candidate materials for further work. They are further used to standardize certain biological products such as vaccines.

The results of acute toxicity tests have a wide variety of regulatory applications. These include determination of the need for childproof packaging, determination of reentry intervals after pesticide application, establishment of the requirement and basis for training workers in chemical use, determination of requirements for protective equipment and clothing, and decision making about general registration of pesticides or their restriction for use by certified applicators. Acute oral toxicity may be used in risk assessments of chemicals for humans and nontarget environmental organisms.

The various national and international regulatory authorities have used different hazard classification systems in the past. In light of the importance of hazard classification, the Organisation for Economic Cooperation and Development (OECD) recently harmonized criteria for hazard classification for global use. For example, the five harmonized categories for acute oral toxicity (in mg/kg body weight) are 0–5, 5–50, 50–300, 300–2000, and 2000–5000.

Acute Oral Testing. Traditionally acute oral toxicity testing has focused on determining the dose that kills half of the animals (i.e., the median lethal dose or LD50), the timing of lethality following acute chemical exposure, as well as observing the onset, nature, severity, and reversibility of toxicity. The LD50 concept was developed by Trevan in 1927. Original testing methods were designed to characterize the dose-response curve by using several animals (usually at least 5/sex) at each of several test doses. Data from a minimum of three doses is required. The LD50 values are presented as estimated doses (mg/kg) with confidence limits. The simplest method for the determination of the LD50 is a graphic one and is based on the assumption that the effect is a

quantal one (all or none), that the percentage responding in an experimental group is dose related, and that the cumulative effect follows a normal distribution. Data from a typical example (see Figure 11.3), their analysis, and implications are discussed in Chapter 11 on acute toxicity.

As a result of much recent controversy, the LD50 test has been the subject of considerable regulatory attention and as a result changes in requirements have been promulgated. These changes are intended to obtain more information but, at the same time, use fewer animals.

Criticism of the LD50 Test. The criticisms of the test include:

- Used uncritically, it is an expression of lethality only, not reflecting other acute effects.
- It requires large numbers of experimental animals to obtain statistically acceptable values. Moreover the results of LD50 tests are known to vary with species, strain, sex, age, and so on (Table 21.4); thus the values are seldom closely similar from one laboratory to another despite the numbers used.
- Because, for regulatory purposes, the most important information needed concerns chronic toxicity, little useful information is derived from the LD50 test. The small amount of information that is acquired could be obtained as well from an approximation requiring only a small number of animals.
- Extrapolation to humans is difficult.

Support of the LD50 Test. Continued use of the test has been advocated, however, on the grounds that it is of use in the following ways:

- Properly conducted, acute toxicity tests yield not only the LD50 but also information on other acute effects such as cause of death, time of death, symptomatology, nonlethal acute effects, organs affected, and reversibility of nonlethal effects.
- Information concerning mode of action and metabolic detoxication can be inferred from the slope of the mortality curve.
- The results can form the basis for the design of subsequent subchronic studies.
- The test is useful as a first approximation of hazards to workers.
- The test is rapidly completed.

For the previously listed reasons, there has been a concerted effort in recent years to modify the concept of acute toxicity testing as it is embodied in the regulations of many countries and to substitute more meaningful methods that use fewer experimental animals. The article by Zbinden and Flury-Roversi is an excellent summary of the

Table 21.4 Factors Causing Variation in LD50 Values

Species	Health	Temperature
Strain	Nutrition	Time of day
Age	Gut contents	Season
Weight	Route of administration	Human error
Gender	Housing	

factors affecting LD50 determinations, the advantages and disadvantages of requiring such tests, and the nature and value of the information derived. It concludes that the acute toxicity test (single-dose toxicity) is still of considerable importance for the assessment of risk posed by new chemical substances, and for a better control of natural and synthetic agents in the human environment. It is not permissible, however, to regard a routine determination of the LD50 in various animal species as a valid substitute for an acute toxicity study.

Current Status. Recently attention has been focused on developing alternatives to the classical LD50 test to reduce the number of animals used or refine procedures to make exposures less stressful to animals. OECD adopted several alternative methods for determining acute oral toxicity: a limit test for materials with anticipated low toxicity, a fixed dose procedure, an acute toxic class method, and an up-and-down procedure. The fixed-dose procedure and the acute toxic class method estimate the LD50 within a dose range for use in classification and labeling. The up-and-down procedure generates point estimates and confidence intervals of the LD50 and therefore may be useful in a wider set of applications.

The fixed-dose procedure (Guideline OECD 420) aims to identify the appropriate hazard class for new chemicals; it does not provide a point estimate of the LD50. This method calls for testing animals sequentially at one of four doses: 5, 50, 300, or 2000 mg/kg body weight. The test begins with a sighting study in which animals are tested, one at a time, at doses selected from the set doses. Once clear signs of toxicity appear, additional animals (females, or the more sensitive sex) are dosed at that level for a total of five animals. Subsequent groups of animals may receive doses at higher or lower levels, if necessary, depending on the outcome of the previous group. Decision criteria based on the number of animals surviving or showing evident toxicity provide for classification decisions.

The acute toxic class method (guideline OECD 423) aims to identify the appropriate hazard and labeling classification and provides a range for lethality rather than a point estimate of the LD50. Groups of three animals (females, or the more sensitive sex) receive one of four or five doses: 5, 50, 300, 2000 and if necessary 5000 mg/kg body weight. Depending on the survival or mortality of the first group of animals, three or more animals may receive the same or a higher or lower dose. The number of animals that survive or die determines the classification decisions.

The up-and-down procedure (Guideline OECD 425) employs sequential dosing, using only a single animal at each step, the dosage depending on whether the previously dosed animal lives or dies. The test provides a point estimate of lethality and confidence intervals, and can be used to evaluate lethality up to 5000 mg/kg. The main test incorporates elements of range finding and uses a flexible stopping point. A sequential limit test uses up to five animals. Default dose spacing is 3.2 times the previous dose. The starting dose should be slightly below the estimated LD50. If no information is available to estimate the LD50, the starting dose is 175 mg/kg. A computer program was developed by the US EPA to simplify both the experimental phase of the test and the calculation of the LD50 and confidence intervals.

For all three guidelines, selection of a starting dose close to the actual LD50 should decrease the number of animals necessary, reduce study duration, and decrease the amount of test substance needed. Therefore it is desirable that all information on the test substance be made available to the testing laboratory for consideration prior to

conducting the study. Such information includes the identity and chemical structure of the substance, its physicochemical properties, the results of any other toxicity test tests on the substance, toxicological data on structurally related substances, the anticipated uses of the substance, or cytotoxicity data on the substance. This information will aid the testing laboratory in selecting the most appropriate test to satisfy regulatory requirements and in choosing the starting dose.

As with the traditional acute oral toxicity methods, the alternative tests involve the administration of a single-bolus dose of a test substance to fasted healthy young adult rodents by oral gavage, observation for morbidity/mortality for up to 14 days after dosing, with recording of body weight (weekly) and clinical signs (daily), and a necropsy at study termination. At the time of dosing, each animal should be between 8 and 12 weeks old and its weight should fall in an interval within $+/-20\%$ of the mean weight of all previously dosed animals taken on their day of dosing. Observation of the postdosing effects on each animal should be for at least 48 hours or until it is clear whether the dosed animals will survive. However, depending on the characteristics of the test material, investigators can vary this time between dosing, so long as the interval is sufficient. Only when the results are clear can a decision be made about whether an additional dose is necessary, and if so, whether to dose the next animal or group of animals at the same, higher, or lower dose. The information from every animal, even those that die after the initial observation period, is used in the final determination of the test outcome.

These newer methods call for testing to be done in a single sex to reduce variability in the test population. This reduction in variability in turn minimizes the number of animals needed. Normally females are used. Although there is usually little difference in sensitivity between males and females, in those cases where there are observable differences, females are most commonly the more sensitive sex. Normally animal suppliers have an excess of female rats because many researchers order only male rats to avoid physiological changes associated with estrus cycling in females; therefore preferential use of female animals for acute testing should not result in excess male animals.

Eye Irritation. Because of the prospect of permanent blindness, ocular toxicity has long been a subject of both interest and concern. Although all regions of the eye are subject to systemic toxicity, usually chronic but sometimes acute, the tests of concern in this section are tests for irritancy of compounds applied topically to the eye. The tests used are all variations of the Draize test, and the preferred experimental animal is the albino rabbit.

The test consists of placing the material to be tested directly into the conjunctival sac of one eye, with the other eye serving as the control. The lids are held together for a few seconds, and the material is left in the eye for at least 24 hours. After that time it may be rinsed out, but in any case, the eye is examined and graded after 1, 2, and 3 days. Grading is subjective and based on the appearance of the cornea, particularly as regards opacity; the iris, as regards both appearance and reaction to light; the conjunctiva, as regards redness and effects on blood vessels; and the eyelids, as regards swelling. Fluorescein dye may be used to assist visual examination because the dye is more readily absorbed by damaged tissues, which then fluoresce when the eye is illuminated. Each end point in the evaluation is scored on a numerical scale, and chemicals are compared on this basis. In addition to the "no-rinse" test, some protocols also investigate the effect of rinsing the eye one minute after exposure to determine

if this reduces the potential for irritation. In addition eyes may be graded for up to 21 days after administration of an irritating test material to evaluate recovery.

The eye irritation test is probably the most criticized by advocates of animal rights and animal welfare, primarily because it is inhumane. It has also been criticized on narrower scientific grounds in that both concentration and volumes used are unrealistically high, and that the results, because of high variability and the greater sensitivity of the rabbit eye, may not be applicable to humans. It is clear, however, that because of great significance of visual impairment, tests for ocular toxicity will continue.

Attempts to solve the dilemma have taken two forms: to find substitute in vitro tests and to modify the Draize test so that it becomes not only more humane but also more predictive for humans. Substitute tests consist of attempts to use cultured cells or eyes from slaughtered food animals, but neither method is yet acceptable as a routine test. Modifications consist primarily of using fewer animals. Usually one animal is tested first and, if the material is severely irritating no further eye testing is conducted. EPA has reduced the required number of animals from 6 to 3. In addition eye irritation should never be carried out on materials with a pH of less than 2 or more than 10 as these materials can be assumed to be potential eye irritants.

Dermal Irritation and Sensitization. There are tests for dermal irritation caused by topical application of chemicals. These fall into four general categories: primary irritation, cutaneous sensitization, phototoxicity, and photosensitization. Because many foreign chemicals come into direct contact with the skin, including cosmetics, detergents, bleaches, and many others, these tests are considered essential to the proper regulation of such products. Less commonly, dermal effects may be caused by systemic toxicants.

In the typical primary irritation test, the backs of albino rabbits are clipped free of hair and an area of about 5 cm^2 on each rabbit is used in the test. This area is then treated with either 0.5 ml or 0.5 g of the compound to be tested and then covered with a gauze pad. The entire trunk of the rabbit is wrapped to prevent ingestion. After 4 to 24 hours the tape and gauze are removed, the treated areas are evaluated for erythematous lesions (redness of the skin produced by congestion of the capillaries) and edematous lesions (accumulation of excess fluid in SC tissue), each of which is expressed on a numerical scale. After an additional 24 to 48 hours, the treated areas are again evaluated.

Skin sensitization tests are designed to test the ability of chemicals to affect the immune system in such a way that a subsequent contact causes a more severe reaction than the first contact. This reaction may be elicited at a much lower concentration and in areas beyond the area of initial contact. The antigen involved is presumed to be formed by the binding of the chemical to body proteins, the ligand-protein complex then being recognized as a foreign protein to which antibodies can be formed. Subsequent exposure may then give rise to an allergic reaction. Skin sensitization tests generally follow protocols that are modifications of the Buchler (dermal inductions) method or the Magnesson and Kligman (intradermal inductions) method. The test animal commonly used in skin sensitization tests is the guinea pig; animals are treated with the test compound in a suitable vehicle, with the vehicle alone, or with a positive control such as 2,4-dinitrochlorobenzene (a relatively strong sensitizer) or cinnamaldehyde (a relatively weak sensitizer) in the same vehicle. During the induction phase the animals are treated for each of 3 days evenly spaced during a 2-week period. This is followed

by a 2-week rest period followed by the challenge phase of the test. This consists of a 24-hour topical treatment carried out as described for primary skin irritation tests. The lesions are scored on the basis of severity and the number of animals responding (incidence). If there is a greater skin reaction in the animals given induction doses compared to those given the test material for the first time, the compound is considered to be a dermal sensitizer.

Other test methods include those in which the induction phase is conducted by intradermal injection together with Freund's adjuvant (a chemical mixture that enhances the antigenic response) and the challenge by dermal application, or tests in which both induction and challenge doses are topical but the former is accompanied by intradermal injections of Freund's adjuvant. It is important that compounds that cause primary skin irritation be tested for skin sensitization at concentrations low enough that the two effects are not confused.

Phototoxicity tests are designed to evaluate the combined dermal effects of light (primarily ultraviolet [UV] light) and the chemical in question. Tests have been developed for both phototoxicity and photoallergy. In both cases the light energy is believed to cause a transient excitation of the toxicant molecule, which, on returning to the lower energy state, generates a reactive, free-radical intermediate. In phototoxicity these organic radicals act directly on the cells to cause lesions, whereas in photoallergy they bind to body proteins. These modified proteins then stimulate the immune system to produce antibodies, because the modifications cause them to be recognized as foreign or "nonself" proteins. These tests are basically modifications of the tests for primary irritation and sensitization except that, following application of the test chemical, the treated area is irradiated with UV light. The differences between the animals treated and irradiated and those treated and not irradiated is a measure of the phototoxic effect.

Safety Pharmacology Studies. Safety pharmacology studies investigate the potential undesirable pharmacodynamic effects of a test article on physiological functions in relationship to exposure. These tests are typically conducted as part of the development of new drugs. The objectives of safety pharmacology studies are threefold: to identify undesirable effects of a test article that may have relevance to its use in humans, to evaluate a test article for possible effects observed in toxicology or clinical studies, and to investigate the mechanism underlying any undesirable effects of the test article. Safety pharmacology consists of a core battery of studies with follow-up studies as indicated by preliminary findings. The core studies are designed to target vital organ systems, particularly the central nervous system, cardiovascular system, and pulmonary system. These studies are typically conducted using small numbers of rats and dogs. In the study for pulmonary function, end points measured are respiratory rate, minute volume, and tidal volume. In the cardiovascular telemetry study, end points include heart rate, blood pressure, and electrocardiogram evaluation. In telemetry studies a radio transmitter is implanted in all animals to permit continuous monitoring for 24 hours pretest and 24 hours after dosing. A cardiopulmonary study can also be conducted in which respiratory rate, minute volume, tidal volume, blood pressure, heart rate, electrocardiogram, and body temperature are monitored in restrained animals for typically 2 hours after dosing.

Subchronic Tests. Subchronic tests examine toxicity caused by repeated dosing over an extended period, but not one that is so long as to constitute a significant portion of

the expected life span of the species tested. A 28- or 90-day oral study in the rat or dog would be typical of this type of study, as would a 21- to 28-day dermal application study or a 28- to 90-day inhalation study. Such tests provide information on essentially all types of chronic toxicity other than carcinogenicity and are usually believed to be essential for establishing the dose regimens for prolonged chronic studies. They are frequently used as the basis for the determination of the no observed effect level (NOEL). This value is often defined as the highest dose level at which no deleterious or abnormal effect can be measured, and is often used in risk assessment calculations. Subchronic tests are also useful in providing information on target organs and on the potential of the test chemical to accumulate in the organism.

Twenty-Eight- to Ninety-Day Tests. Chemicals are usually tested by administration in the diet, less commonly in the drinking water, and only when absolutely necessary by gavage, because the last process involves much handling and subsequent stress. Numerous experimental variables must be controlled and biologic variables evaluated. In addition the number of end points that can be measured is large, and as a consequence record keeping and data analysis must be carefully planned. If all is done with care, much may be learned from such tests.

Experimental (Nonbiologic) Variables. Several environmental variables may affect toxicity evaluations, some directly and others by their effects on animal health. Major deviations from the optimum temperature and humidity for the species in question can cause stress reactions. Stress can also be caused by housing more than one species of experimental animal in the same room. Many toxic or metabolic effects show diurnal variations that are related to photoperiod. Cage design and the nature of the bedding have also been shown to affect the toxic response. Thus the optimum housing conditions are clean rooms, each containing a single species, with the temperature, humidity, and photoperiod being constant and optimized for the species in question. Cages should be the optimum design for the species, bedding should be inert (not cause enzyme induction or other metabolic effect), and cages should not be overcrowded, with individual caging whenever possible.

Dose selection, preparation, and administration are all important variables. Subchronic studies are usually conducted using three (less often, four) dose levels. The highest should produce obvious toxicity but not high mortality and the lowest no toxicity (NOEL), whereas the intermediate dose should give effects clearly intermediate between these two extremes. Although the doses can be extrapolated from acute test, such extrapolation is difficult, particularly in the case of compounds that accumulate in the body, and frequently a 14-day range-finding study is made. Although the route of administration should ideally mimic the expected route of exposure in humans, in practice, the chemical is usually administered ad libitum in the diet because this is, on average, most appropriate. Diets containing known amounts of the test material are presented to the animals. Measurement of food consumption is recommended to provide an estimate of the test material consumed. In cases where a highly accurate measurement of dose is an important factor in the experimental design, the animals may be treated by gavage or by capsules containing the test material.

To avoid effects from nonspecific variations on the diet, enough feed from the same batch should be obtained for the entire study. Part is set aside for the controls, and the remainder is mixed with the test chemical at the various dose levels. Care should

be taken to store all food in such a way that not only does the test chemical remain stable, but also the nutritional value is maintained. The identity and concentration of the test chemical should be checked periodically by chemical analysis. Treated diets may be prepared at set intervals, such as weekly, depending on the stability of the test material in the diet.

Subchronic studies are usually conducted with 10 to 20 males and 10 to 20 females of a rodent species at each dose level and 4 to 8 of each sex of a larger species, such as the dog, at each dose level. Animals should be drawn from a larger group and assigned to control or treatment groups by a random process, but the larger group should not vary so much that the mean weights and ages of the subgroups vary significantly at the beginning of the experiment.

Biologic Variables. Subchronic studies should be conducted on two species, ideally a rodent and a nonrodent. Although the species chosen should be those with the greatest pharmacokinetic and metabolic similarity to humans for the compound in question, this information is seldom available. In practice, the most common rodent used is the rat, and the most common nonrodent used is the dog. It has long been held that inbred rodent strains should be used to reduce variability. This and the search for strains that were sensitive to chemical carcinogenesis but did not have an unacceptably high spontaneous tumor rate led to widespread use of the F344 rat and B6C3F1 mouse. Other researchers believe that an outbred strain such as the Sprague Dawley rat is more robust and prefer to use them.

Although ideally the age should be matched to the expected exposure period in terms of the stage of human development, this is not often done. Young adult or adolescent animals that are still growing are preferred in almost all cases, and both sexes are routinely used.

Good animal care is critical at all times because toxicity has been shown to vary with diet, disease, and environmental factors. Animals should be quarantined for some time before being admitted to the test area, their diet should be optimum for the test species, and the facility should be kept clean at all times. Regular inspection by a veterinarian is essential, and any animals showing unusual symptoms not related to the treatment (e.g., in controls or in low dose but not high dose animals) should be removed from the test and autopsied.

Results. Although the information required from subchronic tests varies somewhat from one regulatory agency to another, the requirements are basically similar (Table 21.5). For explanatory purposes the data obtained from these tests can be described as two types: that obtained from living animals during the course of the test and that obtained from animals sacrificed either during or at the end of the test period. Many of the tests performed on living animals can be carried out first before the test period begins to provide a baseline for comparison to subsequent measurement. A satellite group of treated animals can be added to the test for evaluation of "recovery." For these animals the treated food would be removed at the end of the test period and they would be returned to the control diet for 21 to 28 days while the various end points are followed. This is done to establish whether any effects noted are reversible. Autopsies should be performed on all animals found dead or moribund during the course of the test. The following is a list of end points that may be measured during a 90-day oral toxicity study.

Table 21.5 Summary of Subchronic Test Guidelines by Regulatory Agency

Character of tests	EPA Pesticide Assessment Guidelines	FDA "Red Book"	FDA IND/NDA Pharmacology Review Guidelines	OECD	EPA Health Effects Test Guidelines	NTP
Purpose	Pesticide registration support	Food and color additives; safety assessment	IND/NDR pharmacology review guidelines	Assessment and evaluation of toxic characteristics	Select chronic dose levels	Predict dose range for chronic study
	No-observed-effect level	No observed adverse effects, no-effect level	Characterize pharmacology, toxicology, pharmacokinetics, and metabolism of drugs for precautionary clinical decisions	Select chronic dose levels. Useful information and permissible human exposure	Establish safety criteria for human exposure, no-observed-effect level	
Species	Rat, dog, mouse	Rat, dog	Rat, mouse, other rodents, dog, monkey, other nonrodents	Rat, dog	Rat, dog	Fischer 344 rats $B6C3F_1$ mice
Doses	Three dose levels	Three dose levels	Three dose levels	Three dose levels	Three dose levels	Five dose levels
Endpoints	Clinical signs	Clinical signs				
	Ophthalmology	Ophthalmology	Ophthalmology	Ophthalmology	Ophthalmology	
	Hematology	Hematology	Hematology	Hematology	Hematology	
	Clinical chemistry	Clinical chemistry	Clinical chemistry	Clinical chemistry	Clinical chemistry	
	Histopathology	Histopathology	Histopathology	Histopathology	Histology	Weight loss, histopathology
	Target organs	Target organs	Target organs, behavioral and pharmacological effects	Target organs	Target organs	Target organs

Source: From the National Toxicology/Program, Washington, DC: Department of Health and Human Services. *Report of the NTP Ad hoc Panel on Chemical Carcinogenesis Testing and Evaluation.*
Note: EPA, Environmental Protection Agency; FDA, Food and Drug Administration; IND/NDA, investigative new drug/new drug assessment; OECD, Organization for Economic Cooperation and Development; NTP, National Toxicology Program.

A. In-life Data. Interim tests are carried out at intervals before the study to establish baselines, at intervals during the study, and at the end of the study.

1. *Appearance.* Mortality and morbidity as well as the condition of the skin, fur, mucous membranes, and orifices should be checked at least daily. Presence of palpable masses or external lesions should be noted.
2. *Eyes.* Ophthalmologic examination of both cornea and retina should be carried out at the beginning and at the end of the study.
3. *Food consumption.* Weekly.
4. *Body weight.* Weekly.
5. *Behavioral abnormalities.*
6. *Respiration rate.*
7. *ECG.* Particularly with the larger animals.
8. *EEG.* Particularly with the larger animals.
9. *Hematology.* Assessment should be made prior to chemical administration (pretest) and at least prior to termination. Hemoglobin, hematocrit, RBC, WBC, differential counts, platelets, reticulocytes, and clotting parameters should be assessed.
10. *Blood chemistry.* Pretest, and at least prior to termination, electrolytes and electrolyte balance, acid-base balance, glucose, urea nitrogen, serum lipids, serum proteins (albumin-globulin ratio), enzymes indicative of organ damage such as transaminases and phosphatases should be measured. Toxicant and metabolite levels should be assessed as needed.
11. *Urinalysis.* Pretest, and at least prior to termination, microscopic appearance (sediment, cells, stones, etc.), pH, specific gravity, chemical analysis for reducing sugars, proteins, ketones, and bilirubin should be measured. Toxicant and metabolite levels should be assessed as needed.
12. *Fecal analysis.* Occult blood, fluid content, and toxicant and metabolite levels should be assessed if needed.

B. Termination Tests. Because the number of tissues that may be sampled is large (Table 21.6) and the number of microscopic methods is also large, it is necessary to consider all previous results before carrying out the pathological examination. For example, clinical tests or blood chemistry analyses may implicate a particular target organ that can then be examined in greater detail. All control and high-dose animals are examined in detail. If lesions are found, the next lowest dose group is examined for these lesions, and this method continues until a no effect group is reached.

Because pathology is largely a descriptive science with a complex terminology that varies from one practitioner to another, it is critical that the terminology be defined at the beginning of the study and that the same pathologist examine the slides form both treated and control animals. Pathologist are not in agreement on the necessity or the wisdom of coding slides so that the assessor is not aware of the treatment given the animal from which a particular slide is derived. Such coding however, eliminates unintentional bias, a hazard in a procedure that depends on subjective evaluation. Other items of utmost importance are quality control, slide

Table 21.6 Tissues and Organs to Be Examined Histologically in Chronic and Subchronic Toxicity Tests

Adrenals	Larynx	Salivary gland
Bone and bone marrow	Liver	Sciatic nerve
Brain	Lungs and bronchi	Seminal vesicles
Cartilage	Lymph nodes	Skin
Cecum	Mammary glands	Spinal cord
Colon	Mandibular lymph node	Spleen
Duodenum	Mesenteric lymph node	Stomach
Esophagus	Nasal cavity	Testes
Eyes	Ovaries	Thigh muscle
Gall bladder	Parathyroids	Thymus
Ileum	Pituitary	Urinary bladder
Jejunum	Prostate	Uterus
Kidneys	Rectum	

identification, and data recording. Many tissues may be examined; consequently an even larger number of tissue blocks must be prepared. Because each of these may yield many slides to be stained, comparable quality of staining and the accurate correlation of a particular slide with its parent block, tissue, and animal is critical.

1. *Necropsy.* This must be conducted with care to avoid postmortem damage to the specimens. Tissues are removed, weighed, and examined closely for gross lesions, masses, and so on. Tissues are then fixed in buffered formalin for subsequent histologic examination.
2. *Histology.* The tissues listed in Table 21.6 plus any lesions, masses, or abnormal tissues are embedded, sectioned, and strained for light microscopy. Parafin embedding and staining with hemotoxylin and eosin are the preferred routine methods, but special stains may be used for particular tissues or for a more specific examination of certain lesions. Electron microscopy may also be used for more specific examination of lesions or cellular changes after their initial localization by more routine methods.

Repeated Dose Dermal Tests. Twenty-one to 28-day dermal tests are particularly important when the expected route of human exposure is by contact with the skin, as is the case with many industrial chemicals or pesticides. Compounds to be tested are usually applied daily to clipped areas on the back of the animal, either undiluted or in a suitable vehicle. In the latter case, if a vehicle is used, it is also applied to the controls. Selection of a suitable solvent is difficult because many affect the skin, causing either drying or irritation, whereas others may markedly affect the rate of penetration of the test chemical. Corn oil, methanol, or carboxymethyl cellulose are preferred to dimethyl sulfoxide (DMSO) or acetone. It should also be considered that some of the test chemical may be ingested as a result of grooming by the animal, although this can be controlled to some extent by use of restraining collars and/or wrapping.

The criteria for environment, dose selection, and species selection, for example, are not greatly different from the criteria used for 90-day feeding tests, although the list of end points to be examined is often shorter (e.g., fewer organs may be examined).

It is necessary, however, to pay close attention to the skin at the point of application because local effects may be as important as systemic ones.

Twenty-Eight- to Ninety-Day Inhalation Tests. Inhalation studies are indicated whenever the route of exposure is expected to be through the lungs. Animals are commonly exposed for 6 to 8 hours each day, 5 days each week, in chambers of the type previously discussed. Even in those cases where the animals are maintained in the inhalation chambers during nonexposure hours, food is always removed during exposure. Nevertheless, exposure tends to be in part dermal and, due to grooming of the fur, in part oral. Environmental and biologic parameters are the same as those for other subchronic tests, as are the routine end points to be measured before, during, and after the test period. Particular attention must be paid, however, to effects on the tissues of the nasal cavity and the lungs because these are the areas of maximum exposure.

If the test material is particulate, consideration must be given to the particle size and its inhalation potential. Particles of 4 microns in size are considered to be inhalable; larger particles will be cleared from the respiratory tract by ciliary action and subsequently swallowed (oral exposure) or expelled by sneezing or expectoration.

21.5.2 Chronic Tests

Chronic tests are those conducted over a significant part of the life span of the test animal. The duration of a chronic study is generally one year or more. Typically rat and dog are the preferred species; for carcinogenicity studies, rats and mice are used.

Chronic Toxicity and Carcinogenicity. Descriptions of tests for both chronic toxicity and carcinogenicity are included here because the design is similar—so similar in fact that they can be combined into one test. Chronic toxicity tests are designed to discover any of numerous toxic effects and to define safety margins to be used in the regulation of chemicals. As with subchronic tests, two species are usually used, one of which is either a rat or a mouse strain, in which case the tests are run for 2 years or 1.5 to 2.0 years, respectively. Data are gathered after 1 year to determine chronic effects without potential confounding effects of aging. Data are gathered after 1.5 years (mouse) or 2 years (rat) to determine carcinogenic potential. The nonrodent species used may be the dog, a nonhuman primate, or, rarely, a small carnivore such as the ferret. Chronic toxicity tests may involve administration in the food, in the drinking water, by capsule, or by inhalation, the first being the most common. Gavage is rarely used. The dose used is the maximum tolerated dose (MTD) and usually two lower doses, perhaps 0.25 MTD and 0.125 MTD with the lowest dose being a predicted no effect level.

MTD. The MTD has been defined for testing purposes by the US Environmental Protection Agency (EPA) as:

> [T]he highest dose that causes no more than a 10% weight decrement, as compared to the appropriate control groups; and does not produce mortality, clinical signs of toxicity, or pathologic lesions (other than those that may be related to a neoplastic response) that would be predicted to shorten the animals' natural life span.

This dose is determined by extrapolation from subchronic studies.

The requirements for animal facilities, housing, and environmental conditions are as described for subchronic studies. Special attention must be paid to diet formulation because it is impractical to formulate all of the diets for 2-year study from a single batch. In general, semisynthetic diets of specified components should be formulated regularly and analyzed before use for test material content.

The end points used in these studies are those described for the subchronic study: appearance, ophthalmology, food consumption, body weight, clinical signs, behavioral signs, hematology, blood chemistry, urinalysis, organ weights, and pathology. Some animals may be killed at fixed intervals during the test (e.g., 6, 12, or 18 months) for histologic examination. Particular attention is paid to any organs or tests that showed compound related changes in the subchronic tests.

Carcinogenicity tests have many requirements in common (physical facilities, diets, etc.) with both chronic and subchronic toxicity tests as previously described. Because of the numbers and time required, these tests are usually carried out using rats and/or mice, but in some cases a nonrodent species may also be used. The chemical under test may be administered in the food, in the drinking water, by dermal application, by gavage, or by inhalation, the first two methods being the most common. Because the oncogenic potency of chemicals varies through extreme limits, the purity of the test chemical is of great concern. A 1% contaminant needs only to be 100 times as potent as the test chemical to have an equivalent effect, and differences of this magnitude and greater are not unheard of.

Dosing is carried out over the major part of the life span for rodents, beginning at or shortly after weaning. The highest dose used is the MTD. The principal end point is tumor incidence as determined by histologic examination. The statistical problem of distinguishing between spontaneous tumor occurrence in the controls and chemical-related tumor incidence in the treated animals is great; for that reason large numbers of animals are used. A typical test involves 50 or more rats or mice of each gender in each treatment group. Some animals are necropsied at intermediate stages of the test (e.g., at 12 months), as are all animals found dead or moribund. All surviving animals are necropsied at the end of the test. Tissues to be examined are listed in Table 21.6, with particular attention being paid to abnormal masses and lesions.

21.5.3 Reproductive Toxicity and Teratogenicity

The aim of developmental and reproductive testing is to examine the potential for a compound to interfere with the ability of an organism to reproduce. This includes testing to assess reproductive risk to mature adults as well as the developing individual at various stages of life, from conception to sexual maturity. Traditionally animal studies have been conducted in three "segments": (I) in adults, treatment during a pre-mating period and optionally continuation for the female through implantation or lactation; (II) in pregnant animals treatment during the major period of organogenesis; and (III) treatment of pregnant/lactating animals from the completion of organogenesis through lactation (peri- and postnatal study). Although guidelines addressing treatment regimens have been rather similar throughout the world, required end points measured in adults and developing organisms have varied. International harmonization of guidelines has demonstrated the need for flexibility in testing for reproductive and developmental toxicity, and toxicologists are now often challenged to design unique studies to examine potential effects on all the parameters considered in the classical segment I, II, and III studies. In adults, these include development of mature egg and

sperm, fertilization, implantation, delivery of offspring (parturition), and lactation. In the developing organism, these include early embryonic development, major organ formation, fetal development and growth, and postnatal growth including behavioral assessments and attainment of full reproductive function. These evaluations are usually best carried out in several separate studies.

Some Definitions in Reproductive Biology. Discussion of a bit of reproductive biology may be helpful in the understanding of study designs to evaluate reproductive and developmental toxicity. Tests to assess general reproductive performance and fertility are generally conducted using rats. In the rat, multiple eggs are ovulated from mature follicles in the ovary. The follicle that an egg leaves behind develops into glandular tissue known as a corpora lutea. The corpora lutea secretes progesterone, a hormone needed to maintain pregnancy (unlike humans in which progesterone is secreted by the placenta). Corpora lutea are visible as blisterlike protuberances on the ovary. A count of the corpora lutea in the ovary allows one to determine the maximum number of potential offspring for that pregnancy. Fertilized eggs develop into zygotes that may attach to the wall of the uterus (implantation). The discrete areas of implantation may be observed and counted upon examination of the uterus at C-section. Calculation of pre- and postimplantation loss are important end points in a reproductive toxicity study. Pre-implantation loss is the death of a fertilized ova prior to implantation in the uterine wall. Postimplantation loss (i.e., resorption and/or fetal death) is the death of the conceptus after implantation in the uterine wall and prior to parturition. Postimplantation loss can be broken down into early and late resorptions and fetal death. A late resorption has discernable features such as limbs, eyes, and nose, whereas an early resorption has none of these features.

Single- and Multiple-Generation Tests. Fertility and general reproductive performance can be evaluated in single and multiple generation tests. These tests are usually conducted using rats. Fertility is defined as the ability to produce a pregnancy, while the ability to produce live offspring is known as fecundity. An abbreviated protocol for a single-generation test is shown in Figure 21.2.

In typical tests 25 males per dose group are treated for 70 days prior to mating and 25 females per dose group are treated for 14 days pre-mating. The number of animals is chosen to yield at least 20 pregnant females per dose group including controls. The treatment durations are selected to coincide with critical times during which spermatogenesis and ovulation occur. It takes approximately 70 days in the rat for spermatogonial cells to become mature sperm capable of fertilization. In the female rat the estrus cycle length is 4 to 5 days and a 14-day dosing period is considered sufficient time to detect potential effects on hormonal or other systems that may effect ovulation. In some study designs both males and females are treated for 70 days pre-mating.

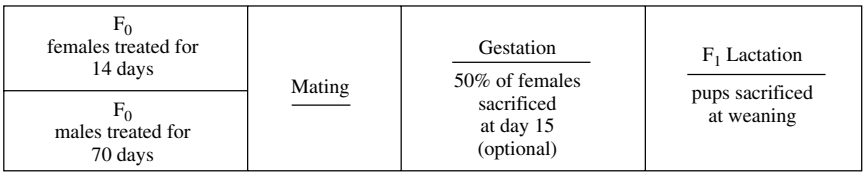

Figure 21.2 Abbreviated protocol for a one generation reproductive toxicity test.

Treatment of the females is continued through pregnancy (21 days) and until the pups are weaned. Pups are usually 21 days of age. The test compound is administered at three dose levels either in the feed, in drinking water, or by gavage. The high dose is chosen to cause some, but not excessive, maternal toxicity (e.g., an approximate 10% decrease in body weight gain, or effects on target organs). Low doses are generally expected to be no-effect levels.

After the pre-mating period, the rats are placed in cohabitation, with one male and one female caged together. Mating is confirmed by the appearance of spermatozoa in a daily vaginal smear. Day 1 of gestation is the day insemination is confirmed. The females bear and nurse their pups. After birth, the pups are counted, weighed, and examined for external abnormalities. The litters are frequently culled to a constant number (usually 8–10) after 4 days. At weaning, the pups are killed and autopsied for gross and internal abnormalities. In a multigeneration study, approximately 25/sex/group are saved to produce the next generation. Brother-sister pairings are avoided. Treatment is continuous throughout the test, which can be carried out for two, sometimes, three generations. An abbreviated protocol for a multiple generation test can be seen in Figure 21.3. Note that the parental generation is known as the F_0 generation and the

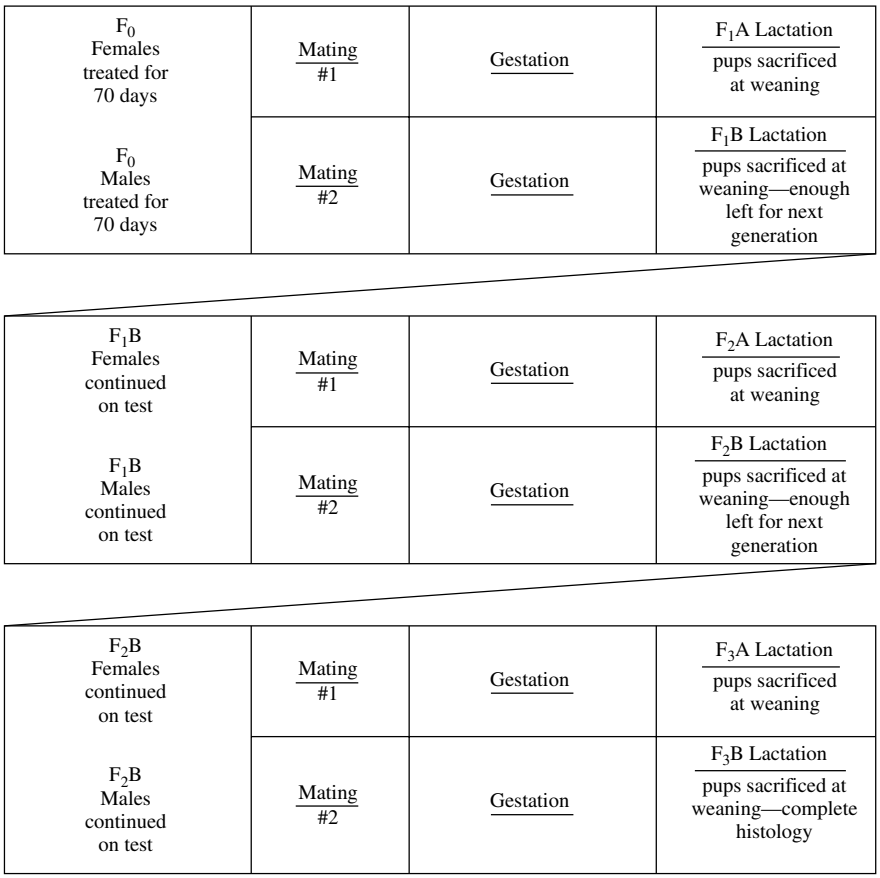

Figure 21.3 Abbreviated protocol for a multigeneration reproductive toxicity test.

offspring are known as the F_1s and F_2s. In some studies, parents produce two litters, for example the F_1A and F_1B litters.

Because both males and females are treated in this type of study design, it is not possible to distinguish between maternal and paternal effects in the reproductive performance. To permit this separation, it is necessary to dose additional animals to the stage of mating and then breed them to untreated members of the opposite sex. Similarly, if effects are seen postnatally, it may not be possible to distinguish between effects mediated in utero or mediated by lactation. This distinction can be made by "cross-fostering" the offspring of treated females to untreated females, and vice versa.

The end points observed in these types of tests, depending on study design, are as follows:

1. Fertility index, the number of pregnancies relative to the number of matings.
2. The number of live births, relative to the number of total births.
3. Pre-implantation loss, or number of corpora lutea in the ovaries relative to the number of implantation sites.
4. Postimplantation loss, or the number of resorption sites in the uterus relative to the number of implantation sites.
5. Duration of gestation.
6. Effects on male or female reproductive systems.
7. Litter size and condition, gross morphology of pups at birth, gender, and anogenital distance.
8. Survival of pups.
9. Weight gain and performance of adults and pups.
10. Time of occurrence of developmental landmarks, such as eye opening, tooth eruption, vaginal opening in females, preputial separation in males.
11. Morphological abnormalities in weanlings.

Results from single and multiple generation tests provide important information for assessment of test materials that may perturb a variety of systems including the endocrine system. A number of variations of the single and multiple generation tests exist. For example, a number of weanlings may be left to develop and be tested later for behavioral and/or physiological defects (e.g., developmental neurotoxicity testing).

Teratology. Teratology is the study of abnormal fetal development. For an agent to be labeled a teratogen, it must significantly increase the occurrence of adverse structural or functional abnormalities in offspring after its administration to the female during pregnancy or directly to the developing organism. In teratology testing, exposure to the test chemical may be from implantation to parturition, although it has also been restricted to the period of major organogenesis, the most sensitive period for inducing structural malformations. Observations may be extended throughout life, but usually they are made immediately prior to birth after a C-section. The end points observed are mainly morphologic (structural changes and malformations), although embryo-fetal mortality is also used as an end point. Figure 21.4 shows an outline of a typical teratology study.

Teratology studies are carried out in two species, a rodent species (usually the rat) and in another species such as the rabbit (rarely in the dog or primate). Enough females

Teratology

Untreated females Untreated males	Mating	Gestation Pregnant females treated on days 6-15. Pups and dams sacrificed day 20

Perinatal/Postnatal

Untreated females Untreated males	Mating	Gestation Pregnant females treated on days 15-21	Lactation Females treated to weaning. Pups and dams sacrificed at weaning

Figure 21.4 Abbreviated protocol for a teratology test and for a perinatal/postnatal toxicity test in rats.

should be used so that, given normal fertility for the strain, there are 20 pregnant females in each dosage group. Traditionally the timing of compound administration has been such that the dam is exposed during the period of major organogenesis, that is days 6 through 15 of gestation in the rat or mouse and days 6 through 18 for the rabbit. Newer study designs call for dosing until C-section. Day 1 is the day spermatozoa appear in the vagina in the case of rats, or the day of mating in the rabbit.

The test chemical is typically administered directly into the stomach by gavage, which is a requirement of EPA and some other regulatory agencies. This method of dosing allows a precise calculation of the amount of test material received by the animal. Studies typically have three dose levels and a control group that receives the vehicle used for test material delivery. The high dose level is chosen to be one at which some maternal toxicity is known to occur, but never one that would cause more than 10% mortality. The low dose should be one at which no maternal toxicity is apparent, and the intermediate dose(s) should be chosen as a predicted low effect level.

The test is terminated by performing a C-section on the day before normal delivery is expected. The uterus is examined for implantation and resorption sites and for live and dead fetuses, and the ovaries are examined for corpora lutea. In rodent studies, half of the fetuses are examined for soft tissue malformations, and the remaining are examined for skeletal malformations. In nonrodents, all fetuses are examined for both soft tissue and skeletal malformations. The various end points that may be examined include maternal toxicity, embryo-fetal toxicity, external malformations, and soft tissue and skeletal malformations.

Careful evaluation of maternal toxicity is necessary in assessing the validity of the high-dose level, and the possibility that maternal toxicity is involved in subsequent events. The parameters evaluated include body weight, food consumption, clinical signs, and necropsy data such as organ weights. Because exposure starts after implantation, conception and implantation rates should be the same in controls and all treatment levels. If not, the test is suspect, with a possible error in the timing of the dose or use of animals from a source unsuitable for this type of testing.

Embryo-fetal toxicity is determined from the number of dead fetuses and resorption sites relative to the number of implantation sites. In addition to the possibility of lethal malformations, such toxicity can be due to maternal toxicity, stress, or direct toxicity

to the embryo or fetus that is not related to developmental malformations. Fetal weight and fetal size may also be a measure of toxicity but should not be confused with the variations seen as a result of differences in the number of pups per litter. Smaller litters tend to have larger pups while larger litters have smaller pups.

Anomalies may be regarded either as variations that may not adversely affect the fetus and not have a fetal outcome or as malformations that have adverse effects on the fetus. For some findings there is disagreement as to which class it belongs, such as the number of ribs in the rabbit that inherently has a large amount of variability. Common external anomalities are listed in Table 21.7 and are determined by examination of

Table 21.7 External Malformations Commonly Seen in Teratogenicity Tests

Brain, cranium, spinal cord
Encephalocele—protrusion of brain through an opening of the skull.
Cerebrum is well formed and covered by transparent connective tissue.
Exencephaly—lack of skull with disorganized outward growth of the brain.
Microcephaly—small head on normal sized body.
Hydrocephaly—marked enlargement of the ventricles of the cerebrum.
Craniorachischisis—exposed brain and spinal cord.
Spina bifida—Nonfusion of spinal processes. Usually ectoderm covering is missing and spinal cord is evident.
Nose
Enlarged naris—enlarged nasal cavities
Single naris—a single naris, usually median
Eye
Microphthalmia—small eye
Anophthalmia—lack of eye
Open eye—no apparent eyelid, eye is open
Ear
Anotia—absence of the external ear
Microtia—small ear
Jaw
Micrognathia—small lower jaw
Agnathia—absence of lower jaw
Aglossia—absence of tongue
Astomia—lack of mouth opening
Bifid tongue—forked tongue
Cleft lip—either unilateral or bilateral cleft of upper lip
Palate
Cleft palate—a cleft or separation of the median portion of the palate
Limbs
Clubfoot—foot that has grown in a twisted manner, resulting in an abnormal shape or position. It is possible to have a malposition of the whole limb
Micromelia—abnormal shortness of the limb
Hemimelia—absence of any of the long bones, resulting in a shortened limb
Phomelia—absence of all of the long bones of a limb, the limb is attached directly to the body

Table 21.8 Some Common Visceral Anomalies Seen in Teratogenicity Tests

Intestines
Umbilical hernia—protrusion of the intestines into the umbilical cord
Ectopic intestines—extrusion of the intestines outside the body wall
Heart
Dextrocardia—rotation of the heart axis to the right
Enlarged heart—either the atrium or the ventricle may be enlarged
Lung
Enlarged lung—all lobes are usually enlarged
Small lung—all lobes are usually small; lung may appear immature
Uterus/testes
Undescended testes—testes are located anterior to the bladder instead of lateral; may be bilateral or unilateral
Agenesis of testes—one or both testes may be missing
Agenesis of uterus—one or both horns of the uterus may be missing
Kidney
Hydronephrosis—fluid-filled kidney, often grossly enlarged; may be accompanied by a hydroureter (enlarged, fluid-filled ureter)
Fused—kidneys fused, appearing as one misshapen kidney with two ureters
Agenesis—one or both kidneys missing
Misshapen—small, enlarged (usually internally), or odd-shaped kidneys

fetuses at C-section. Visceral anomalities are determined by examination of fetuses after fixation using either the dissection method of Staples or by the hand-sectioning method of Wilson. Common visceral findings are listed in Table 21.8. Fetal skeletons are examined after first fixing the fetus and then staining the bone with Alizarin red. Numerous skeletal variations occur in controls and may not have an adverse effect on the fetus (Table 21.9). Their frequency of occurrence may, however, be dose related and should be evaluated.

Almost all chemically induced malformations have been observed in control animals, and most malformations are known to be produced by more than one cause. Thus it is obvious that great care is necessary in the interpretation of teratology studies. For an agent to be classified as a development toxicant or teratogen, it must produce adverse effects on the conceptus at exposure levels that do not induce toxicity in the mother. Signs of maternal toxicity include reduction in weight gain, changes in eating patterns, hypo or hyperactivity, neurotoxic signs, and organ weight changes. Adverse effects on development under these conditions may be secondary to stress on the maternal system. Findings in the fetus, at dose levels that produce maternal toxicity, can not be easily separated from the maternal toxicity. Compounds can be deliberately administered at maternally toxic dose levels to determine the threshold for adverse effects on the offspring. In such cases conclusions can be qualified to indicate that adverse effects on the offspring were found at maternally toxic dose levels and may not be indicative of selective or unique developmental toxicity.

Effect of Chemicals in Late Pregnancy and Lactation (Perinatal and Postnatal effects). These tests are usually carried out on rats, and 20 pregnant females per

Table 21.9 Skeletal Abnormalities Commonly Seen in Teratogenicity Tests

Digits
Polydactyly—presence of extra digits, in mouse six or more, instead of five
Syndactyly—fusion of two or more digits
Oligodactyly—absence of one or more digits
Brachydactyly—smallness of one or more digits
Ribs
Wavy—ribs may be any aberrant shape
Extra—may have extra ribs on either side
Fused—may be fused anywhere along the length of the rib
Branched—single base and branched
Tail
Short—short tail, usually lack of vertebrae
Missing—absence of tail
Corkscrew—corkscrew-shaped tail

dosage group are treated during the final third of gestation and through lactation to weaning (day 15 of pregnancy through day 21 postpartum) (Figure 21.4). The duration of gestation, parturition problems, and the number and size of pups in the naturally delivered litter are observed, as is the growth performance or the offspring. Variations of this test are the inclusion of groups treated only to parturition and only postpartum in order to separate prenatal and postnatal effects. Cross-fostering of pups to untreated dams may also be used to the same end. Behavioral testing of the pups has been suggested, and this and other physiological testing is to be recommended.

21.5.4 Special Tests

This general heading is used to include brief assessments of tests that are not always required but that may be required in certain cases or that have been suggested as useful adjuncts to current testing protocols. Many are in areas of toxicology that are developing rapidly; as a result no consensus has yet evolved as to the best tests or sequence of tests, only an understanding that such evaluations may shed light on previously undefined aspects of chemical toxicity.

Neurotoxicity. The nervous system is complex, both structurally and functionally, and toxicants can affect one or more units of this system in selective fashion. It is necessary therefore to devise tests, or sequences of tests, that measure not only changes in overall function but that also indicate which basic unit is affected and how the toxicant interacts with its target. This is complicated by the fact that the nervous system has a considerable functional reserve, and specific observable damage may not affect overall function until it becomes even more extensive. Types of damage to the nervous system are classified in various ways but include neuronal toxicity, axonopathy, toxic interruption of impulse transmission, myelinopathy, and synaptic alterations in transmitter release or receptor function. Signs of neuropathy are frequently revealed by the acute, subchronic, chronic, and other tests that are required by regulatory agencies.

Neurotoxicity is of great significance in toxicology, however, and tests have been devised to supplement those routinely required. These include acute and subchronic neurotoxicity studies as well as developmental neurotoxicity studies.

Behavioral and Pharmacological Tests. Behavioral and pharmacological tests involve the observation of clinical signs and behavior. These include signs of changes in awareness, mood, motor activity, central nervous system excitation, posture, motor incoordination, muscle tone, reflexes, and autonomic functions. If these tests so indicate, more specialized tests can be carried out that evaluate spontaneous motor activity, conditioned avoidance responses, operant conditioning, as well as tests for motor incoordination such as the inclined plane or rotarod tests.

Tests for specific classes of chemicals include the measurement of transmitter stimulated adenyl cyclase and Na/K-ATPase for chemicals that affect receptor function or cholinesterase inhibition for organophosphates or carbamates. Electrophysiological techniques may detect chemicals such as DDT or pyrethroids, which affect impulse transmission.

Acute and Subchronic Neurotoxicity Tests. These studies are designed to test a wide range of neurotoxicity effects including CNS stimulation or depression, reflex perturbation, peripheral nerve damage, cognitive effects on learning and memory, motor activity effects and neuropathology. The tests are conducted using rats or sometimes mice.

In the acute neurotoxicity study, approximately 10 to 15 animals per sex per dose group are administered a single gavage (bolus) dose of the test material. There are usually three dose groups and a control. Behavioral assessments are made on the day of dosing, and at 1 and 2 weeks post dose. The assessments include tests on motor activity, a functional observation battery (FOB), and neuropathology (at termination). The FOB screens for sensorimotor, neuromuscular, autonomic and general physiological effects of a test compound. Table 21.10 depicts component tests of the FOB. These functional tests have the advantage over biochemical measures that they permit repeated evaluation of individual animals over time to determine the onset, progression, duration, and reversibility of neurotoxic effects. Motor activity is also measured over time and can be evaluated by a variety of devices. One such device that has been frequently used is the figure 8 maze, which consists of a series of interconnected alleys (in the shape of the numeral 8) converging on a central open area and covered with transparent acrylic plastic. Motor activity is detected by photobeams, and an activity count is registered each time a photobeam is interrupted by the animal. Motor activity sessions are generally 60 minutes in length and each session is divided into 5- to 10-minute reporting intervals (epochs). "Habituation" is an end point evaluated in the motor activity test, and this is defined as a decrement in activity during the test session. Activity is expected to decrease toward the end of the test as the animal's exploratory activity normally lessens as the time in the maze increases. Neuropathological examinations are the same as those described below for the subchronic neurotoxicity test.

Before the acute neurotoxicity study is conducted, it is necessary to conduct a preliminary test to determine the time of peak effect after dosing of the test material. Preliminary tests may evaluate a selected group of end points in the FOB or other sensitive end points, if known, for a particular test material. Results of this preliminary test will determine the time when observations are performed on the day of dosing in the acute neurotoxicity study.

Table 21.10 Example of Evaluations Made in a Functional Observation Battery

Home Cage and Open Field	Manipulative	Physiological
Arousal	Ease of removal	Body temperature
Gait	Ease of handling	Body weight
Posture	Touch response	
Vocalizations	Righting response	
Piloerection	Hindlimb foot splay	
Lacrimation	Forelimb grip strength	
Salivation	Hindlimb grip strength	
Urination/defecation	Finger-snap response	
Grooming behavior	Catalepsy	
Rearing	Palpebral closure	
Abnormal movements	Pupil function	
Tremors, convulsions		

In the subchronic neurotoxicity study, end points measured are similar to those measured in the acute neurotoxicity study. However, the duration of dosing is 90 days and exposure to the test material is usually via the diet. As for the acute neurotoxicity study, these studies consist of three test groups and a control group. The functional observation battery and motor activity tests are conducted at selected intervals such as weeks 5, 9, and 13, as well as pretest. At test termination, at least 6 animals per group are perfused via the heart with fixative to ensure optimal fixation of nervous tissues for histopathology examination. Nervous tissues examined include brain, spinal cord (various segments), and selected nerves such as the optic, sciatic, tibial, and sural nerves.

For all behavioral tests it is important that the person making the actual observations is unaware of the treatment group for each animal ("blind" to dose group assignment). In addition laboratories that conduct neurotoxicity studies for regulatory agencies must demonstrate that their methods are validated. Therefore these laboratories must conduct positive control studies using known neurotoxicants and provide this information to regulators as necessary. Also, since it is not feasible for one person to perform the observations on all animals on all test occasions, laboratories must maintain evidence of interobserver reliability (agreement) for individuals who are involved with performing the functional observation batteries.

Developmental Neurotoxicity Tests. A separate component of developmental toxicology focuses on potential behavioral or morphological modifications resulting from exposures to toxins during early development. These studies track the outcome of such exposures through the postnatal period and into early adulthood. In a developmental neurotoxicity study at least 20 pregnant female rats for each of three treatment groups plus a control are administered test material from gestation day (GD) 6 through weaning on lactation day 21. FOBs are conducted on the maternal animals at selected intervals such as GD 6, GD 17, lactation day 11 and lactation day 21. Evaluations include observations in the home cage, during handling, and outside the home cage in an open field. Body weights and food consumption are also monitored in the maternal animals. After birth the offspring are counted, weighed, and gender determined. On

postnatal day (PND) 4, litters are culled to 8 pups per litter. Following culling, at least 10 pups/sex/group are assigned to one of the following tests: learning and memory, motor activity, or acoustic startle. Additional pups are assigned for neuropathology and brain weight evaluations on PND 11 and PND 70 (10/sex/group). FOBs are performed on the offspring at selected intervals such as PND 11, 17, 21, 35, and 60/70. Indicators of physical development such as preputial separation (male sexual maturation) and vaginal patency (female sexual maturation) are evaluated for all offspring as well as body weight and food consumption. Learning and memory can be evaluated with a variety of tests. Frequently a water maze is used where the rat learns to swim through a series of alleys to find a platform it can use to climb out of the water. The time it takes for the animal to swim through the maze to the platform (trial latency) and the number of mistakes made are some of the end points evaluated in this test. The trials are conducted over a series of days and the assay provides an index of the development of both working memory (with-in day performance) as well as reference memory (between-day performance). The startle test measures the animal's response to a burst of loud noise and also how quickly it becomes habituated to 10 pulses of startle-eliciting tones in 5 blocks. For the startle test special chambers lined with sound-attenuating and vibration-absorbing material are used. These chambers can measure the force exerted on a platform on which the animal stands during the test procedure. The startle test is conducted at 2 time points such as PND 22 and 60. When the offspring are approximately 70 days old, the test is terminated. Selected animals are perfused with fixative for neuropathology assessments. In addition to detailed microscopic evaluation of at least five different sections of the brain, simple morphometric analysis of the cerebrum, hippocampus, and cerebellum are conducted.

Delayed Neuropathy (OPIDN). The delayed neurotoxic potential of certain organophosphates such as tri-o-cresyl phosphate (TOCP) is usually evaluated by observation of clinical signs (paralysis of leg muscles in hens) or pathology (degeneration of the motor nerves in hens), but a biochemical test involving the ratio between the ability to inhibit cholinesterase and the ability to inhibit an enzyme that has been referred to as the neurotoxic esterase (NTE) has been suggested. The ability of chemicals to cause delayed neuropathy is generally correlated with their ability to inhibit this nonspecific esterase, found in various tissues, although the role, if any, of NTE in the sequence of events leading to nerve degeneration is not known. The preferred test organism is the mature hen because the clinical signs are similar to those in humans, and such symptoms cannot be readily elicited in the common laboratory rodents.

Potentiation. Potentiation and synergism represent interactions between toxicants that are potential sources of hazard because neither humans nor other species are usually exposed to one chemical at a time. The enormous number of possible combinations of chemicals makes routine screening for all such effects not only impractical but impossible.

One of the classic cases is the potentiation of the insecticide malathion by another insecticide, EPN, the LD50 of the mixture being dramatically lower than that of either compound alone. This potentiation can also be seen between malathion and certain contaminants that are formed during synthesis, such as isomalathion. For this reason quality control during manufacture is essential. This example of potentiation involves inhibition, by EPN or isomalathion, of the carboxylesterase responsible for the detoxication of malathion in mammals.

It is practical to test for potentiation only when there has been some preliminary indication that it might occur or when either or both compounds belong to chemical classes previously known to cause potentiation. Such a test can be conducted by comparing the LD50, or any other appropriate toxic end point, of a mixture of equitoxic doses of the chemicals in question with the same end point measured with the two chemicals administered alone.

In the case of synergism, in which one of the compounds is relatively nontoxic when given alone, the toxicity of the toxic compound can be measured when administered alone or after a relatively large dose of the nontoxic compound.

Toxicokinetics and Metabolism. Routine toxicity testing without regard to the mechanisms involved is likely to be wasteful of time and of human, animal, and financial resources. A knowledge of toxicokinetics and metabolism can give valuable insights and provide for testing that is both more efficient and more informative. Such knowledge provides the necessary background to make the most appropriate selection of test animal species and of dose levels, and the most appropriate method for extrapolating from animal studies to the assessment of human hazard. Moreover they may provide information on possible reactive intermediates as well as information on induction or inhibition of the enzymes of xenobiotic metabolism, the latter being critical to an assessment of possible interaction.

The nature of metabolic reactions and their variations between species is detailed in Chapters 7, 8, and 9 with some aspects of toxicokinetics in Chapter 6. The methods used for the measurement of toxicants and their metabolites are detailed in Chapter 25. The present section is concerned with the general principles, use, and need for metabolic and toxicokinetics studies in toxicity testing.

Toxicokinetics studies are designed to measure the amount and rate of the absorption, distribution, metabolism, and excretion of a xenobiotic. These data are used to construct predictive mathematical models so that the distribution and excretion of other doses can be simulated. Such studies are carried out using radiolabeled compounds to facilitate measurement and total recovery of the administered dose. This can be done entirely in vivo by measuring levels in blood, expired air, feces, and urine; these procedures can be done relatively noninvasively and continuously in the same animal. Tissue levels can be measured by sequential killing and analysis of organ levels. It is important to measure not only the compound administered but also its metabolites, because simple radioactivity counting does not differentiate among them.

The metabolic study, considered separately, consists of treatment of the animal with the radiolabeled compound followed by chemical analysis of all metabolites formed in vivo and excreted via the lungs, kidneys, or bile. Although reactive intermediates are unlikely to be isolated, the chemical structure of the end products may provide vital clues to the nature of the intermediates involved in their formation. The use of tissue homogenates, subcellular fractions, and purified enzymes may serve to clarify events occurring during metabolic sequences leading to the end products.

Information of importance in test animal selection is the similarity in toxicodynamics and metabolism to that of humans. Although all of the necessary information may not be available for humans, it can often be inferred with reference to metabolism and excretion of related compounds, but it is clearly ill advised to use an animal that differs from most others in the toxicokinetics or metabolism of the compound in question or that differs from humans in the nature of the end products. Dose selection

is influenced by knowledge of whether a particular dose saturates a physiological process such as excretion or whether it is likely to accumulate in a particular tissue, because these factors are likely to become increasingly important the longer a chronic study continues.

Behavior. Although the primary emphasis in toxicity testing has long been the estimation of morphologic changes, much recent interest has focused on more fundamental evaluations. One such aspect has been the evaluation of chemical effects on behavior.

The categories of methods used in behavioral toxicology fall into two principal classes, stimulus-oriented behavior, and internally generated behavior. The former includes two types of conditioned behavior: operant conditioning, in which animals are trained to perform a task in order to obtain a reward or to avoid a punishment, and classical conditioning, in which an animal learns to associate a conditioning stimulus with a reflex action. Stimulus-oriented behavior also involves unconditioned responses in which the animal's response to a particular stimulus is recorded.

Internally generated behavior includes observation of animal behavior in response to various experimental situations, and includes exploratory behavior, circadian activity, social behavior, and so on. The performance of animals treated with a particular chemical is compared with that of untreated controls as a measure of the effect of the chemical.

Many of the variables associated with other types of testing must also be controlled in behavioral tests: sex, age, species, environment, diet, and animal husbandry. Behavior may vary with all of these. Norton describes a series of four tests that may form an appropriate series, inasmuch as they represent four different types of behavior; the series should therefore reflect different types of nervous system activity. They are as follows:

1. *Passive avoidance.* This test involves the use of a shuttle box, in which animals can move between a light side and a dark side. After an acclimatization period, in which the animal can move freely between the two sides, it receives a mild electric shock while in the dark (preferred) side. During subsequent trials, the time spent in the "safe side" is recorded.
2. *Auditory startle.* This test involves the response (movement) to a sound stimulus either without, or preceded by, a light-flash stimulus.
3. *Residential maze.* Movements of animals in a residential maze are automatically recorded during both light and dark photoperiods.
4. *Walking patterns.* Gait is measured in walking animals, including such characteristics as the length and width of stride and the angles formed by the placement of the feet.

Problems associated with behavioral toxicology include the functional reserve and adaptability of the nervous system. Frequently behavior is maintained despite clearly observable injury. Other problems are the statistical ones associated with multiple tests, multiple measurements, and the inherently large variability in behavior.

The use of human subjects occupationally exposed to chemicals is often attempted, but such tests are complicated by the subjective nature of the end points (dizziness, etc.).

Covalent Binding. Toxicity has been associated with covalent binding in a number of ways. Organ-specific toxicants administered in vivo bind covalently to macromolecules, usually at a higher level in the target tissues than in nontarget tissues. Examples include acetaminophen in the liver, carbon tetrachloride in the liver, *p*-aminophenol in the kidney, and ipomeanol in the lung. Similarly many carcinogens are known to give rise to DNA adducts. In general, covalent binding occurs as a result of metabolism of the toxicant to highly reactive intermediates, usually, but not always, by cytochrome P450. Because these intermediates are highly reactive electrophiles, they bind to many nucleophilic sites on DNA, RNA, or protein molecules, not just the site of toxic action. Thus measurement of covalent binding may be a measure of toxic potential rather than a specific measurement, related directly to a mechanism of action. The occurrence of covalent binding at the same time as toxicity is so common an occurrence, however, that a measurement of covalent binding of a chemical may be regarded as an excellent although perhaps not infallible indication of potential for toxicity. Although such tests are not routine, considerable interest has been shown in their development.

The measurement of DNA adducts is an indirect indication of genotoxic (carcinogenic) potential, and DNA adducts in the urine are an indication, obtained by a noninvasive technique, of recent exposure. Protein adducts give an integrated measure of exposure because they accumulate over the life span of the protein and, at the same time, indicate possible organ toxicity.

Tissue protein adducts are usually demonstrated in experimental animals following injection of radiolabeled chemicals and, after a period to time, the organs are removed and homogenized, and by rigorous extraction, all the noncovalently bound material is removed. Extraction methods include lipid solvents, acids and bases, concentrated urea solutions, and solubilization and precipitation of the proteins. They tend to underestimate the extent of covalent binding because even covalent bonds may be broken by the rigorous procedures used. Newer methods involving dialysis against detergents and separation of adducted proteins will probably prove more appropriate.

Blood proteins, such as hemoglobin, may be used in tests of human exposure because blood is readily and safely accessible. For example, the exposure of mice to ethylene oxide or dimethylnitrosamine was estimated by measuring alkylated residues in hemoglobin. The method was subsequently extended to people exposed occupationally to ethylene oxide by measuring *N*-3-(2-hydroxyethyl) histidine residues in hemoglobin. Similarly methyl cysteine residues in hemoglobin can be used as a measure of methylation.

DNA-RNA adducts can also be measured in various ways, including rigorous extraction, separation, and precipitation following administration of labeled compounds in vivo, or use of antibodies raised to chemically modified DNA or RNA.

Although many compounds of different chemical classes have been shown to bind covalently when activated by microsomal preparations in vitro (e.g., aflatoxin, ipomeanol, stilbene, vinyl chloride), these observations have not been developed into routine testing procedures. Such procedures could be useful in predicting toxic potential.

Immunotoxicity. Immunotoxicology comprises two distinct types of toxic effects: the involvement of the immune system in mediating the toxic effect of a chemical and the toxic effects of chemicals on the immune system. The former is shown, for

example, in tests for cutaneous sensitization, whereas the latter is shown in impairment of the ability to resist infection.

Tests for immunotoxicity are not required by all regulatory agencies, but it is an area of great interest, both in the fundamental mechanisms of immune function and in the design of tests to measure impairment of immune function. Both of these aspects are discussed in detail in Chapter 19.

21.6 IN VITRO AND OTHER SHORT-TERM TESTS

21.6.1 Introduction

The toxicity tests that follow are tests conducted largely in vitro with isolated cell systems. Some are short-term tests carried out in vivo or are combinations of in vivo and in vitro systems. The latter are included because of similarities in approach, mechanism, or intent. In general, these tests measure effects on the genome or cell transformation; their importance lies in the relationship between such effects and the mechanism of chemical carcinogenesis. Mutagenicity of cells in the germ line is itself an expression of toxicity, however, and the mutant genes can be inherited and expressed in the next or subsequent generations.

The theory that the initiating step of chemical carcinogenesis is a somatic mutation is well recognized, and considerable evidence shows that mutagenic potential is correlated with carcinogenic potential. Thus the intent of much of this type of testing is to provide early warning of carcinogenic potential without the delay involved in conducting lifetime chronic feeding studies in experimental animals. Despite the numerous tests that have been devised, regulatory agencies have not yet seen it fit to substitute any of them, or any combination of them, for chronic feeding studies. Instead, they have been added as additional testing requirements. One function of such tests should be to identify those compounds with the greatest potential for toxicity and enable the amount of chronic testing to be reduced to more manageable proportions.

21.6.2 Prokaryote Mutagenicity

Ames Test. The Ames test, developed by Bruce Ames and his coworkers at the University of California, Berkeley, depends on the ability of mutagenic chemicals to bring about reverse mutations in *Salmonella typhimurium* strains that have defects in the histidine biosynthesis pathway. These strains will not grow in the absence of histidine but can be caused to mutate back to the wild type, which can synthesize histidine and hence can grow in its absence. The postmitochondrial supernatant (S-9 fraction), obtained from homogenates of livers of rats previously treated with PCBs in order to induce certain cytochrome P450 isoforms, is also included in order to provide the activating enzymes involved in the production of the potent electrophiles often involved in the toxicity of chemicals to animals.

Bacterial tester strains have been developed that can test for either base-pair (e.g., strain TA-1531) or frameshift (e.g., strains TA-1537, TA-1538) mutations. Other, more sensitive strains such as TA-98 and TA-100 are also used, although they may be less specific with regard to the type of mutation caused.

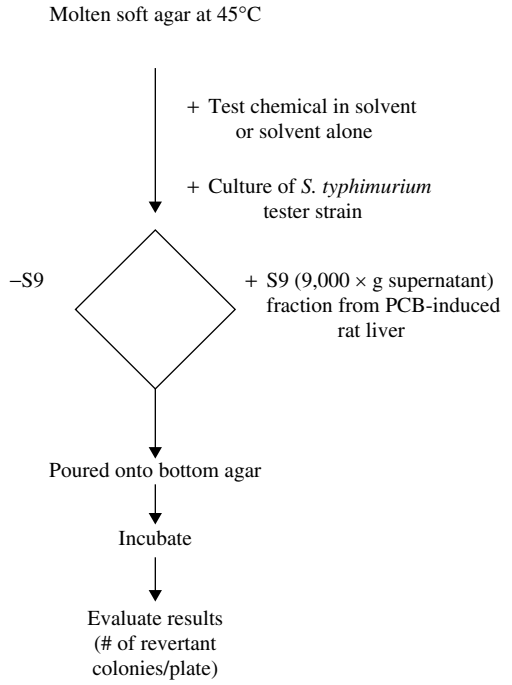

Figure 21.5 Protocol for the Ames test for mutagenicity.

In brief, the test is carried out (Figure 21.5) by mixing a suspension of bacterial cells with molten top agar. This also contains cofactors, S-9 fraction, and the material to be tested. The mixture is poured onto Petri plates containing hardened minimal agar. The number of bacteria that revert and acquire the wild-type ability to grow in the absence of histidine can be estimated by counting the colonies that develop on incubation. To provide a valid test, a number of concentrations are tested, and positive controls with known mutagens are included along with negative controls that lack only the test compound. The entire test is replicated often enough to satisfy appropriate statistical tests for significance. Parallel tests without the S-9 fraction may help distinguish between chemicals with intrinsic mutagenic potential and those that require metabolic activation.

The question of correlation between mutagenicity and carcinogenicity is crucial in any consideration of the utility of this or similar tests. In general, this appears to be high, although a small proportion of both false positives and false negatives occurs. For example, certain base analogues and inorganics such as manganese are not carcinogens but are mutagens in the Ames test, whereas diethylstilbestrol (DES) is a carcinogen but not a bacterial mutagen (see Chapter 12 for additional detail).

Related Tests. Related tests include tests based on reverse mutations, as in the Ames test, as well as tests based on forward mutations. Examples include:

1. *Reverse mutations in Escherichia coli.* This test is similar to the Ames test and depends on reversion of tryptophane mutants, which cannot synthesize this amino acid, to the wild type, which can. The S-9 fraction from the liver of induced rats

can also be used as an activating system in this test. Other *E. coli* reverse mutation tests utilize nicotinic acid and arginine mutants.

2. *Forward mutations in S. typhimurium.* One such assay, dependent on the appearance of a mutation conferring resistance to 8-azaguanine in a histidine revertant strain, has been developed and is said to be as sensitive as the reverse-mutation tests
3. *Forward mutations in E. coli.* These mutations depend on mutation of galactose nonfermenting *E. coli* to galactose fermenting *E. coli* or the change from 5-methyltryptophane to 5-methyltryptophane resistance.
4. *DNA repair.* Polymerase-deficient, and thus DNA repair-deficient, *E. coli* has provided the basis for a test that depends on the fact that the growth of a deficient strain is inhibited more by a DNA-damaging agent than is that of a repair-competent strain. The recombinant assay using *Bacillus subtilis* is conducted in much the same way because recombinant deficient strains are more sensitive to DNA-damaging agents.

21.6.3 Eukaryote Mutagenicity

Mammalian Cell Mutation. The development of cell culture techniques that permit both survival and replication have led to many advances in cell biology, including the use of certain of these cell lines for detection of mutagens. Although such cells, if derived from mammals, would seem ideal for testing for toxicity toward mammals, there are several problems. Primary cells, which generally resemble those of the tissue of origin, are difficult to culture and have poor cloning ability. Because of these difficulties, certain established cell lines are usually used. These cells, such as Chinese hamster ovary cells and mouse lymphoma cells, clone readily and do not become senescent with passage through many cell generations. Unfortunately, they have little metabolic activity toward xenobiotics and thus do not readily activate toxicants. Moreover they usually show chromosome changes, such as aneuploidy (i.e., more or fewer than the usual diploid number of chromosomes).

The characteristics usually involved in these assays are resistance to 8-azaguanine or 6-thioguanine (the hypoxanthine guanine phosphoribosyl transferase or HGPRT locus), resistance to bromodeoxyuridine or triflurothymidine (the thymidine kinase or TK locus), or resistance to ouabain (the OU or Na/K-ATPase locus). HGPRT is responsible for incorporation of purines from the medium into the nucleic acid synthesis pathway. Its loss prevents uptake of normal purines and also of toxic purines such as 8-azaguanine, which would kill the cell. Thus mutation at this locus confers resistance to these toxic purine analogues. Similarly TK permits pyrimidine transport, and its loss prevents uptake of toxic pyrimidine analogues and confers resistance to them. In the absence of HGPRT or TK, the cells can grow by de novo synthesis of purines and pyrimidines. Ouabain kills cells by combining with the Na/K-ATPase. Mutation at the OU locus alters the ouabain-binding site in a way that prevents inhibition and thus confers resistance.

A typical test system is the analysis of the TK locus in mouse lymphoma cells for mutations that confer resistance to bromodeoxyuracil. The tests are conducted with and without the S-9 fraction from induced rat liver because the lymphoma cells have little activating ability. Both positive and negative controls are included, and the parameter

measured is the number of cells formed that are capable of forming colonies in the presence of bromodeoxyuridine.

***Drosophila* Sex-Linked Recessive Lethal Test.** The advantages of *Drosophila* (fruit fly) tests are that they involve an intact eukaryotic organism with all of its interrelated organ systems and activation mechanisms but, at the same time, are fast, relatively easy to perform, and do not involve mammals as test animals. The most obvious disadvantages are that the hormonal and immune systems of insects are significantly different from those of mammals and that the nature, specificity, and inducibility of the cytochrome P450s are not as well understood in insects as they are in mammals.

In a typical test, males that are 2 days postpuparium and that were raised from eggs laid within a short time period (usually 24 hours) are treated with the test compound in water to which sucrose has been added to increase palatability. Males from a strain carrying a gene for yellow body on the X chromosome are used. Preliminary tests determine that the number of offspring of the survivors of the treatment doses (usually 0.25 LD50 and 0.5 LD50) are adequate for future crosses. Appropriate controls, including a solvent control (with emulsifier if one was necessary to prepare the test solution), and a positive control, such as ethyl methane sulfonate, are routinely included with each test. Individual crosses of each surviving treated male with a series of three females are made on a 0- to 2-, 3- to 5-, and 6- to 8-day schedule. The progeny of each female is reared separately, and the males and females of the F_1 generation are mated in brother-sister matings. If there are no males with yellow bodies in a particular set of progeny, it should be assumed that a lethal mutation was present on the treated X chromosomes. A comparison of the F_2 progeny derived from females inseminated by males at different times after treatment allows a distinction to be made between effects on spermatozoa, spermatids, and spermatocytes.

In the Basc (Muller-5) test shown in Figure 21.6, the strain used for the females in the F_1 cross is a multiple-marked strain that carries a dominant gene for bar eyes and recessive genes for apricot eyes and a reduction of bristles on the thorax (scute gene). (Basc is an acronym for bar, apricot, and scute.)

Related Tests. Many tests related to the two types of eukaryote-mutation tests are discussed earlier in this section, and many of them are simply variations of the tests described. Two distinct classes are worthy of mention: the first uses yeasts as the test organisms, and the second is the spot test for mutations in mice.

One group of yeast tests includes tests for gene mutations and strains that can be used to detect forward mutations in genes that code for enzymes in the purine biosynthetic pathway; other strains can be used to detect reversions. Yeasts can also be used to test for recombinant events such as reciprocal mitotic recombination (mitotic crossing over) and nonreciprocal mitotic recombination. Saccharomyces cerevisiae is the preferred organism in almost all these tests. Although they possess cytochrome P450s capable of metabolizing xenobiotics, their specificity and sensitivity are limited as compared with those of mammals, and an S-9 fraction is often included, as in the Ames test, to enhance activation.

The gene mutation test systems in mice include the specific locus test, in which wild-type treated males are crossed with females carrying recessive mutations for visible phenotypic effects. The F_1 progeny have the same phenotype as the wild-type parent unless a mutation, corresponding to a recessive mutant marker, has occurred. Such tests

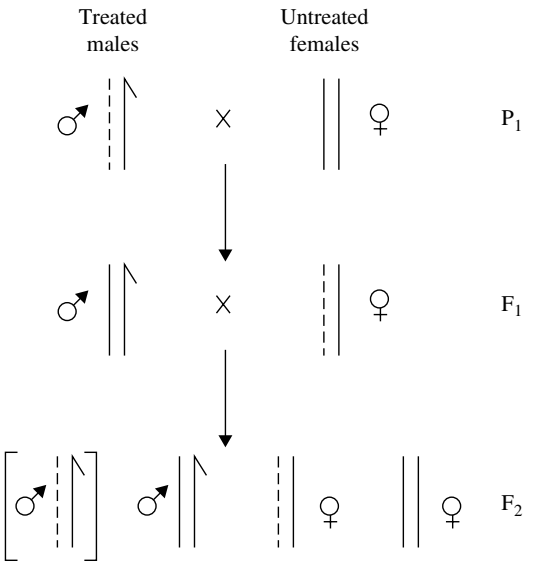

Figure 21.6 The Basc (Muller-5) mating scheme. Dashed lines represent the treated X chromosome of males. Brackets indicate males with yellow bodies, which would be absent if a lethal mutation occurred on the X chromosome of the treated male.

are accurate, and the spontaneous (background) mutation rate is very low, making them sound tests that are predictive for other mammals. Unfortunately, the large number of animals required has prevented extensive use. Similar tests involving the activity and electrophoretic mobility of various enzymes in the blood or other tissues in the F_1 progeny from treated males and untreated females have been developed. In the previously mentioned tests, as with many others, sequential mating of males with different females can provide information about the stage of sperm development at which the mutational event occurred.

21.6.4 DNA Damage and Repair

Many of the end points for tests described in this chapter, including gene mutation, chromosome damage, and oncogenicity, develop as a consequence of damage to or chemical modification of DNA. Most of these tests, however, also involve metabolic events that occur both prior to and subsequent to the modification of DNA. Some tests, however, use events at the DNA level as end points. One of these, the unscheduled synthesis of DNA in mammalian cells, is described in some detail; the others are summarized briefly.

Unscheduled DNA Synthesis in Mammalian Cells. The principle of this test is that it measures the repair that follows DNA damage and is thus a reflection of the damage itself. It depends on the autoradiographic measurement of the incorporation of tritiated thymidine into the nuclei of cells previously treated with the test chemical.

The preferred cells are usually primary hepatocytes in cultures derived from adult male rats whose cells are dispersed and allowed to attach themselves to glass coverslips.

From this point on, the test is carried out on the attached cells. Both positive controls with agents known to stimulate unscheduled DNA synthesis, such as the carcinogen aflatoxin B1 or 2-acetylaminofluorene, and negative controls, which are processed through all procedures except exposure to the test compound, are performed routinely with every test. Cells are exposed by replacing the medium for a short time with one containing the test chemical. The dose levels are determined by a preliminary cell viability test (Trypan blue exclusion test) and consist of several concentrations that span the range from no apparent loss of viability to almost complete loss of viability. Following exposure, the medium is removed, and the cells are washed by several changes of fresh medium and finally placed in a medium containing tritiated thymidine. The cells are fixed and dried, and the coverslip with the cells attached is coated with photographic emulsion. After a suitable exposure period (usually several weeks), the emulsion is developed and the cells are stained with hemotoxylin and eosin. The number of grains in the nuclear region is corrected by subtracting nonnuclear grains, and the net grain count in the nuclear area is compared between treated and untreated cells.

This test has several advantages in that primary liver cells have considerable activation capacity and the test measures an event at the DNA level. It does not, however, distinguish between error-free repair and error-prone repair, the latter being itself a mutagenic process. Thus it cannot distinguish between events that might lead to toxic sequelae and those that do not. A modification of this test measures in vivo unscheduled DNA synthesis. In this modification animals are first treated in vivo, and primary hepatocytes are then prepared and treated as already described.

Related Tests. Tests for the measurement of binding of the test material to DNA have already been discussed under covalent binding (Section 21.5.4). Another method of assessing DNA damage is the estimation of DNA breakage following exposure to the test chemical; the DNA-strand length is estimated by using alkaline elution or sucrose density gradient centrifugation. This has been done with a number of cell lines and with freshly prepared hepatocytes, in the latter case following treatment either in vivo or in vitro. It may be regarded as promising but not yet fully validated. The polymerase-deficient *E. coli* tests as well as recombinant tests using yeasts are also related to DNA repair.

21.6.5 Chromosome Aberrations

Tests for chromosome aberrations involve the estimation of effects on extended regions of whole chromosomes rather than on single or small numbers of genes. Primarily they concern chromosome breaks and the exchange of material between chromosomes.

Sister Chromatid Exchange. Sister chromatid exchange (SCE) occurs between the sister chromatids that together make up a chromosome. It occurs at the same locus in each chromatid and is thus a symmetrical exchange of chromosome material. In this regard it is not strictly an aberration because the products do not differ in morphology from normal chromosome. SCE, however, is susceptible to chemical induction and appears to be correlated with the genotoxic potential of chemicals as well as with their oncogenic potential. The exchange is visualized by permitting the treated cells to pass through two DNA replication cycles in the presence of 5-bromo-2′-deoxyuridine, which

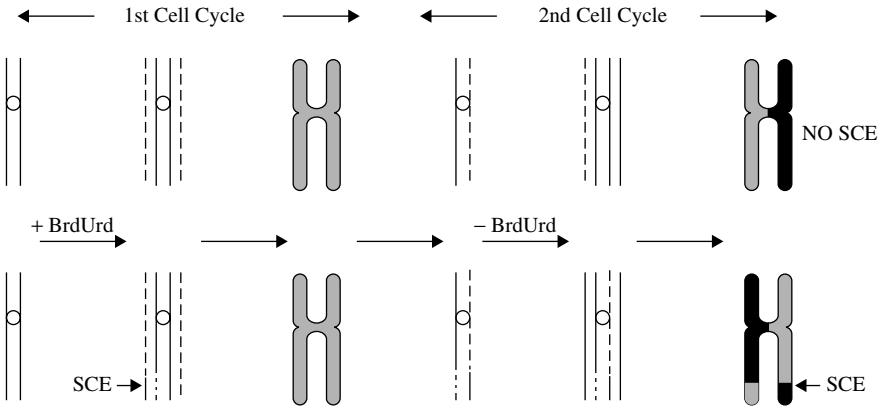

Figure 21.7 Visualization of sister chromatid exchange.

is incorporated in the replicated DNA. The cells are then stained with a fluorescent dye and irradiated with UV light, which permits differentiation between chromatids that contain bromodeoxyuridine and those that do not (Figure 21.7).

The test can be carried out on cultured cells or on cells from animals treated in vivo. In the former case the test chemical is usually evaluated in the presence and absence of the S-9 activation system from rat liver. Typically cells from a Chinese hamster ovary cell line are incubated in a liquid medium and exposed to several concentration of the test chemical, either with or without the S-9 fraction, for about 2 hours. Positive controls, such as ethyl methane sulfonate (a direct-acting compound) or dimethylnitrosamine (one that requires activation), as well as negative controls are also included. Test concentrations are based on cell toxicity levels determined by prior experiment and are selected in such a way that even at the highest dose excess growth does not occur. At the end of the treatment period the cells are washed, bromodeoxyuridine is added, and the cells are incubated for 24 hours or more. The cells are then fixed, stained with a fluorescent dye, and irradiated with UV light. Second division cells are scored under the microscope for SCEs (Figure 21.7).

The test can also be carried out on cells treated in vivo, and analyses have been made of SCEs in lymphocytes from cancer patients treated with chemotherapeutic drugs, smokers, and workers exposed occupationally; in several cases increased incidence of SCEs has been noted. This is a sensitive test for compounds that alkylate DNA, with few false positives. It may be useful for detecting promoters such as phorbol esters.

Micronucleus Test. The micronucleus test is an in vivo test usually carried out in mice. The animals are treated in vivo, and the erythrocyte stem cells from the bone marrow are stained and examined for micronuclei. Micronuclei represent chromosome fragments or chromosomes left behind at anaphase. It is basically a test for compounds that cause chromosome breaks (clastogenic agents) and compounds that interfere with normal mitotic cell division, including compounds that affect spindle fiber function.

Male and female mice from an outbred strain are handled by the best animal husbandry techniques, as described for acute, subchronic, and chronic tests, and are treated either with the solvent, 0.5 LD50, or 0.1 LD50 of the test chemical. Animals are killed at several time intervals up to 2 days; the bone marrow is extracted, placed on

microscope slides, dried, and stained. The presence of micronuclei is scored visually under the microscope.

Dominant Lethal Test in Rodents. The dominant lethal test, which is performed using rats, mice, or hamsters, is an in vivo test to determine the germ-cell risk from a suspected mutagen. The test consists of treating males with the test compound for several days, followed by mating to different females each week for enough weeks to cover the period required for a complete spermatogenic cycle. Animals are maintained under optimal conditions of animal husbandry and are dosed, usually by a gavage, with several doses of less than 0.1 LD50. The females are killed after two weeks of gestation and dissected; corpora lutea and living and dead implantations are counted. The end points used to determine the occurrence of dominant lethal mutations in the treated males are the fertility index (ratio of pregnant females to mated females), preimplantation losses (the number of implantations relative to the number of corpora lutea), the number of females with dead implantations relative to the total number of pregnant females, and the number of dead implantations relative to the total number of implantations. Mutations in sperm that are dominant and lethal do not result in viable offspring.

Related Tests. Many cells exposed to test chemicals can be scored for chromosome aberrations by staining procedures followed by visual examination with the aid of the microscope. These include Chinese hamster ovary cells in culture treated in a protocol very similar to that used in the test for SCEs, bone marrow cells from animals treated in vivo, or lymphocytes from animals treated in vivo. The types of aberrations evaluated include chromatid gaps, breaks, and deletions; chromosome gaps, breaks, and deletions; chromosome fragments; translocations; and ploidy.

Heritable translocations can be detected by direct examination of cells from male or female offspring in various stages of development or by crossing the treated animals to untreated animals and evaluating fertility, with males with reduced fertility being examined for translocations, and so on. Progeny from this or other tests, such as those for dominant lethals, can be permitted to survive and then examined for translocations and other abnormalities.

21.6.6 Mammalian Cell Transformation

Most cell transformation assays utilize fibroblast cultures derived from embryonic tissue. The original studies showed that cells from C3H mouse fibroblast cultures developed morphologic changes and changes in growth patterns when treated with carcinogens. Later similar studies were made with Syrian hamster embryo cells. The direct relationship of these changes to carcinogenesis was demonstrated by transplantation of the cells into a host animal and the subsequent development of tumors. The recent development of practical assay procedures involves two cell lines from mouse embryos, Balb/3T3 and C3H/10T1/2, in which transformation is easily recognized and scored. In a typical assay situation, cells, such as Balb/3T3 mouse fibroblasts, will multiply in culture until a monolayer is formed. At this point they cease dividing unless transformed. Chemicals that are transforming agents will, however, cause growth to occur in thicker layers above the monolayer. These clumps of transformed cells are known as foci. Despite many recommended controls the assay is only semiquantitative.

The doses are selected from the results of a preliminary experiment and range from a high dose that reduces colony formation (but not by >50%) to a low dose that has no measurable effect on colony formation. After exposure to the test chemical for 1 to 3 days, the cells are washed and incubation is continued for up to 4 weeks. At that time the monolayers are fixed, stained, and scored for transformed foci.

Transformation assays have several distinct advantages. Because transplanted foci give rise to tumors in congenic hosts (those from the same inbred strain from which the cells were derived) whereas untransformed cells do not, cell transformation is believed to be illustrative of the overall expression of carcinogenesis in mammalian tissues. The two cell types used most (Balb/3T3 and C3H/10T1/2) respond to promoters in the manner predicted by the multistage model for carcinogenesis in vivo and may eventually be useful in the development of assays for promotion. Unfortunately, a large number of false negative results are obtained because these cell lines do not show much activation capacity; it has not proved practical to combine them with the S-9 activation system. Furthermore the cells are aneuploidy and may be preneoplastic in the untreated state. Syrian hamster cells, which do have considerable activation capacity, have proved difficult to use in test procedures and are difficult to score.

21.6.7 General Considerations and Testing Sequences

Considering all of the tests for acute and chronic toxicity, long and short term, in vivo and in vitro, it is clearly impractical to apply a complete series of tests to all commercial chemicals and all their derivatives in food, water, and the environment. The challenge of toxicity testing is to identify the most effective set or sequence of tests necessary to describe the apparent and potential toxicity of a particular chemical or mixture of chemicals. The enormous emphasis on in vitro or short-term tests that has occurred since the mid-1970s had its roots in the need to find substitutes for lifetime feeding studies in experimental animals or, at the very least, to suggest a sequence of tests that would enable priorities to be set for which chemicals should be subjected to chronic tests. Such tests might also be used to eliminate the need for chronic testing for chemicals that either clearly possessed the potential for toxicity or clearly do not. Although there has been much success in test development, the challenge outlined here has not been met, primarily because of the failure of scientists and regulatory agencies, worldwide, to agree on test sequences or on the circumstances in which short-term tests may substitute for chronic tests. Thus not only are short-term tests often required, these tests are in addition to long-term tests. As an example, the US EPA requirements for the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) include, besides a full battery of acute, subchronic, and chronic tests, tests to address the following three categories: gene mutations, structural chromosome aberrations, and other genotoxic tests as appropriate (e.g., DNA damage and repair and chromosome aberrations). It is important, however, that test sequences have been suggested and considered by regulatory agencies, but there must also be taken into account the fact that short-term tests do not provide all of the information needed from the longer term tests.

21.7 ECOLOGICAL EFFECTS

Tests for ecological effects include those designed to address the potential of chemicals to affect ecosystems and the population dynamics in the environment. The tests

are conducted to estimate effects on field populations of vertebrates, invertebrates, and plants. The use of environmental risk assessment tests is discussed in detail in Chapter 28.

21.7.1 Laboratory Tests

There are two types of laboratory tests: toxicity determinations on wildlife and aquatic organisms and the use of model ecosystems to measure bioaccumulation and transport of toxicants and their degradation products.

Among the tests included in the first category are the avian oral LD50, the avian dietary LC50, wild mammal toxicity, and avian reproduction. The avian tests are usually carried out on bobwhite quail or mallard ducks, whereas the wild mammals may be species such as the pine mouse, *Paramyscus*. The tests are similar to those described under acute and chronic testing procedures but suffer from some drawbacks; the standards of animal husbandry used with rats and mice are probably unattainable with birds or wild mammals, even through bobwhite quail and mallards are easily reared in captivity. The genetics of the birds and mammals used are much more variable than are those of the traditional laboratory rodent strains.

Similar tests can be carried out with aquatic organisms (e.g., the LC50 for freshwater fish such as rainbow trout and bluegills), the LC50 for estuarine and marine organisms, the LC50 for invertebrates such as *Daphnia*, and the effect of chemicals on the early stages of fish and various invertebrates.

Model systems, first developed by ecologists to study basic ecological processes, have been adapted to toxicological testing. In toxicology these models were first used to determine the movement and concentration of pesticides. Typically the model has a water phase containing vertebrates and invertebrates, and a terrestrial phase containing at least one plant species and one herbivore species. First, the ^{14}C-labeled pesticide or other environmental contaminant is applied to the leaves of the terrestrial plant sorghum *(Sorgum halpense)*, and then salt marsh caterpillars (*Estigmene acrea*) are placed on the plants. The larvae eat the plants and contaminate the water with feces and their dead bodies. The aquatic food chain is simulated with plankton (diatoms, rotifers, etc.), water fleas (*Daphnia*), mosquito larvae (*Culex pipiens*), and fish (*Gambusia affinis*). From an analysis of the plants, animals, and substrates for the ^{14}C-labeled compound and its degradation products, the biologic magnification or rate of degradation can be calculated.

More complex models involving several compartments, simulated rain, simulated soil drainage, simulated tidal flow, and so on, have been constructed and their properties investigated, but none have been brought to the stage of use in routine testing. Similarly aquatic models using static, re-circulating, and continuous flow have also been used, as have entirely terrestrial models: again, none have been developed for routine testing.

21.7.2 Simulated Field Tests

Simulated field tests may be quite simple, consisting of feeding treated prey to predators and studying the toxic effects on the predator, enabling some predictions concerning effects to nontarget organisms. In general, however, the term is used for greenhouse, small plot, small artificial pond, or small natural pond tests. These serve to test biologic

accumulation and degradation under conditions somewhat more natural than in model ecosystems and the test chemicals are exposed to environmental as well as biologic degradation. Population effects may be noted, but these methods are more useful for soil invertebrates, plants, and aquatic organisms because other organisms are not easily contained in small plots.

21.7.3 Field Tests

In field-test situations, test chemicals are applied to large areas under natural conditions. The areas are at least several acres and may be either natural or part of some agroecosystem. Because the area is large and in the open, radiolabeled compounds cannot be used, it is not possible to obtain a balance between material applied and material recovered.

The effects are followed over a long period of time and two types of control may be used: first, a comparison with a similar area that is untreated; and second, and a comparison with the same area before treatment. In the first case it is difficult, if not impossible, to duplicate exactly a large natural area, and in the second, changes can occur that are unrelated to the test material.

In either case, studies of populations are the most important focus of this type of testing, although the disappearance of the test material, its accumulation in various life forms, and the appearance, accumulation, and disappearance of its degradation products are also important. The population of soil organisms, terrestrial organisms, and aquatic organisms as well as plants all must be surveyed and characterized, both qualitatively and quantitatively. After application of the test material the populations can be followed through two or more annual cycles to determine both acute and long-term population effects.

21.8 RISK ANALYSIS

The preceding tests for various kinds of toxicity can be used to measure adverse effects of many different chemical compounds in different species, organ, tissues, cells, or even populations, and under many different conditions. This information can be used to predict possible toxicity of related compounds from QSAR or of the same chemical under different conditions (e.g., mutagenicity as a predictor of carcinogenicity). It is considerably more difficult to use this information to predict possible risk to other species, such as humans, because little experimental data on this species is available. Some methods are available to predict risk to humans and to provide the risk factor in the risk-benefit assessment that provides the basis for regulatory action, however. Human health risk assessment is discussed in detail in Chapter 24. The benefit factor is largely economic in nature, and the final regulatory action is not, in the narrow sense, a scientific one. It also involves political and legal aspects and, in toto, represents society's evaluation of the amount of risk that can be tolerated in any particular case.

21.9 THE FUTURE OF TOXICITY TESTING

Because of the public awareness of the potentially harmful effects of chemicals, it is clear that toxicity testing will continue to be an important activity and that it will be

required by regulatory agencies before the use of a particular chemical is permitted either in commercial processes or for use by the public. Because of the proliferation of testing procedures, the number of experimental species and other test systems available, as well as the high dose rates usually used, it is clear that eventually some expression of some type of toxicity will be obtained for most exogenous chemicals. Thus the identification of toxic effects with the intent of banning any chemical causing such effects is no longer a productive mode of attack. The aim of toxicity testing should be to identify those compounds that present an unacceptable potential for risk to humans or to the environment and thus ought to be banned but, at the same time, provide an accurate assessment of the risk to humans and the environment of less toxic compounds so that their use may be regulated.

Subjecting all chemicals to all possible tests is logistically impossible, and the future of toxicity testing must lie in the development of techniques that will narrow the testing process so that highly toxic and relatively nontoxic compounds can be identified early and either banned or permitted unrestricted use without undue waste of time, funds, and human resources. These vital commodities could then be concentrated on compounds whose fate and effects are less predictable.

Such progress will come from further development and validation of the newer testing procedures and the development of techniques to select, for any given chemical, the most suitable testing methods. Perhaps of most importance is the development of integrated test sequences that permit decisions to be made at each step, thereby either abbreviating the sequence or making the next step more effective and efficient.

SUGGESTED READING

Balls, M., R. J. Riddell, and A. N. Worden. *Animals and Alternatives in Toxicity Testing*. London: Academic Press, 1983.

Clark, B., and D. A. Smith. Pharmacokinetics and toxicity testing. *CRC Crit. Rev. Toxicol.* **12**: 343, 1984.

Couch, J. A., and W. J. Hargis Jr. Aquatic animals in toxicity testing. *J. Am. Coll. Toxicol.* **3**: 331, 1984.

Dean, J. H., M. I. Luster, M. J. Murray, and L. D. Laver. Approaches and methodology for examining the immunological effects of xenobiotics. *Immunotoxicol.* **7**: 205, 1983

de Serres F. J., Ashby J. (eds): *Evaluation of Short Term Tests for Carcinogens*. New York: Elsevier, 1981.

Ecobichon, D. J.. *The Basis of Toxicity Testing*. Boca Raton, FL: CRC Press, 1992.

Enslein, K., and P. N. Craig. Carcinogenesis: A predictive structure-activity model. *J. Toxicol. Environ. Health* **10**: 521, 1982.

Gorrod, J. W., ed. *Testing for Toxicity*. London: Taylor and Francis, 1981. Relevant chapters include:

Chapter 3. Brown, V. K. H.: Acute toxicity testing—a critique.

Chapter 4. Roe, F. J. C.: Testing in vivo for general chronic toxicity and carcinogenicity.

Chapter 7. Gorrod, J. W.: Covalent binding as an indication of drug toxicity.

Chapter 15. Dewar, A. J.: Neurotoxicity testing—with particular reference to biochemical methods.

Chapter 18. Cobb, L. M.: Pulmonary toxicity.

Chapter 20. Parish, W. E.: Immunological tests to predict toxicological hazards to man.

Chapter 21. Venitt, S.: Microbial tests in carcinogenesis studies.

Chapter 22. Styles, J. A.: Other short-term tests in carcinogenesis studies.

Hayes, A. W., ed. *Principles and Methods of Toxicology*, 2nd ed. New York: Raven Press. This volume contains the following chapters of particular relevance to this chapter:

Chapter 6. Chan, P. K., and A. W. Hayes: Principles and methods for acute and eye irritancy.

Chapter 7. Mosberg, A. T., and A. W. Hayes: Subchronic toxicity testing.

Chapter 8. Stevens, K. R., and M. A. Gallo: Practical considerations in the conduct of chronic toxicity studies.

Chapter 9. Roberts, J. F., W. W. Piegorsch, and R. L. Schueler: Methods in testing for carcinogenicity.

Chapter 10. Zenick, H., and E. D. Clegg: Assessment of male reproductive toxicology: a risk assessment approach.

Chapter 11. Manson, J. M., and Y. J. Kang: Test methods for assessing female reproductive and developmental toxicology.

Chapter 12. Kennedy, G. L., Jr: Inhalation toxicology.

Chapter 13. Patrick, E., and H. I. Maiback: Dermatotoxicology.

Chapter 14. Brusick, D.: Genetic toxicology.

Chapter 17. Burger, G. T., and L. C. Miller: Animal care and facilities.

Chapter 18. Norton, S.: Methods for behavioral toxicology.

Chapter 26. Dean, J. H., et al.: Immune system; evaluation of injury.

Chapter 30. Renwick, A. G.: Pharmacokinetics in toxicology.

Chapter 31. Hogan, M. D., and D. G. Hoel: Extrapolation to humans.

Moser, V. C., G. C. Becking, V. Cuomo, et al. The IPCS collaborative study on neurobehavioral screening methods: IV. Control data. *Neurotoxicol.* **18**: 947–967, 1997.

Weiss, B., and D. Cory-Slechta. Assessment of behavioral toxicity. In *Principles and Methods of Toxicology*, 3rd ed. A. Wallace Hayes, ed. New York: Raven Press, 1994.