Appendix C

Methods for Toxicologic Testing *

by SRI International

METHODS FOR TOXICOLOGIC TESTING

In evaluating substances for their toxicologic effects, several essentials must be a part of each test protocol so that the toxicologic profile is as accurate as possible within the limitations of each test. This is true whether the test material is used for *in vitro* assays or administered *in vivo* by inhalation, gavage, dietary, dermal, subcutaneous, intravenous, or any other route to experimental animals. Both the testing and subsequent evaluation of data must be done under the direction of individuals having the necessary education, training, and experience to conduct the testing and interpretation of results in accordance with sound scientific principles and good laboratory practices. These principles and practices have been addressed and elaborated on in the FDA proposed guidelines (1) for "Good Laboratory Practices in Non-Clinical Laboratory Studies." It is important to realize that, at present, extrapolations from animal test systems to humans to predict effects of exposure to hazardous materials can be made only tentatively.

The data used for the evaluation of potential risk should be derived from tests conducted with the form of the toxicant that poses the health hazard. In some cases, this will be the technical grade of the active ingredient. In other cases, it could be the manufactured end product, the pure grade of the material, an inert ingredient of the test substance, or a metabolite or degradation product of the substance so that the study can be completed using the same lot sample. Within the limits of analytical detectability, the chemical composition of the test material must be determined. If the test substance is to be mixed with the diet or another vehicle, the concentration and homogeneity must be ascertained prior to beginning the study. During the course of the investigation, random samples should be taken to ensure that the required concentration of the substance is maintained.

Healthy animals must be used and maintained under conditions conforming to good husbandry practices (1). Animals must be assigned to test groups in a way that minimizes bias and assures that the number of animals in each dose group and the number of dose groups are sufficient to yield statistically valid results. The animals used in the study should be of uniform weight and age and should be an adequate representation of the sex, species, and strain under consideration. In addition to the treated groups, negative control groups (both historic and concurrent) meeting the above specifications must also be used to evaluate any toxic manifestations that may have occurred independent of the test substance. In certain situations, positive control groups may be required. This is particularly necessary for many short-term *in vitro* assays. Concurrent controls must be handled in the same way as the treated animals so that the treated and control groups will be validly comparable. Control groups take on added importance if a carrier is used in administering the test material to confirm that the carrier selected for use is nontoxic, induces no independent physiological effects, and does not change the chemical characteristics or toxicity of the test material.

Toxicology testing methods can be categorized by duration and endpoint. Short-term tests are usually considered to require 90 days or less for data gathering. These may include simple tests, such as 2-hour LD50 range finding, or more complex ones, such as 90-day continuous exposure or paired feeding studies. Tests requiring more than 90 days, such as lifetime exposure studies, are generally considered long-term tests. In addition to the time necessary for exposure and data gathering, analysis of the results may take up to an additional year depending on the complexity of the experiment, the number of animals used in the experiment, and the volume of data collected. Methods may also be categorized by endpoint. In this

*Excerpt from OTA Working Paper entitled "Assessment of Methods for Regulating 'Unavoidable' Contaminants in the Food Supply." A complete copy of the paper can be obtained from the National Technical Information Service. (See app. J.)

scheme, experiments are designed and data are collected based on expected results such as functional systemic changes, teratogenicity, or carcinogenicity. By the use of an appropriate experimental design, several endpoints can be assessed in the same experimental period such as is done in FDA's three-generation studies (2).

For ease of presentation, this appendix has been subdivided by endpoint into sections on Systemic Toxicity, Carcinogenicity, Mutagenicity, Teratology and Effects on Reproduction, Metabolism, and Structure-Activity Relationships. This appendix is not intended as an exhaustive survey of all testing methods used, but is meant to give an overview of those methods most commonly used today by toxicologists.

Local and Systemic Toxicity

Some of the fastest and simplest methods for determining the toxicities of substances involve the observation of changes in the structure and function of organs and organ systems. These methods generally involve absolute and relative weight changes, gross and microscopic structural alterations, and primary and secondary tests for organ, system, or whole animal function. With advances in the chemical, physiological, and behavioral sciences, modifications for testing systemic toxicity have been proposed that make these procedures more sophisticated and relatively complicated. Several good texts are available which review systemic toxicity [3,4).

Range Finding

The classic determinations of toxicity involve percent lethal or effective dose, concentration, or time. These tests may employ any route of exposure, the ones chosen usually being based on factors such as chemical and physical properties of the agent and potential routes of exposure from the environment. The results obtained from these determinations are usually specific for the species, sex, age, and condition of the organism, and for the route of exposure and environmental conditions before, during, and after exposure. The endpoints of these tests may be either structural or functional changes, but they are usually limited to gross effects such as death or narcosis. These tests are primarily used to determine relative toxicities of various agents and for range finding for maximum tolerated dosage preliminary to beginning a subacute study. They are not usually used to directly evaluate the hazard.

In general, these tests will employ young adult rats and another mammalian nonrodent species. Selection of this other species " . . should consider such factors as comparative metabolism of the chemical and species sensitivity to the toxic effects of the test substance . . . . "(l) The route of administration chosen is that most nearly identical to the potential human exposure. Doses are usually chosen to give results in the 20- to 80-percent lethal or effective range and are usually separated by 0.5 log units (5). Many modifications of this basic procedure are accepted.

Irritation

The irritation potential of substances is tested by observation of the reflex behavior of the animal and by direct observation of the site of contact with the agent, Attempts to quantitate reflex behavioral responses, i.e., eye rubbing, regurgitation, or shallow breathing, have met with little success. The simplest protocols for evaluating irritation involve the skin and eyes. Semiquantitative systems for scoring skin and eye irritation have been proposed by several authors (6-12) and involve placing the suspected irritant in contact with the skin or eye of New Zealand White rabbits. Protocols for skin irritation involve contact with both intact and abraded skin to differentiate the agent's ability to penetrate the skin barrier, and occlusion of the contact site to maximize the response. The severity of erythema and edema is scored as the endpoint. Eye irritation studies are carried out without washing and with washing at various intervals to determine the effectiveness of removal of the agent to reducing the adverse effect, Opacity, area affected, iris reaction to light, hemorrhage, swelling, and discharge are scored. Results from these tests can vary greatly depending on the method of application of the substance, whether dry or premoistened, etc.

Other potential sites of irritation such as the sensory nerves, respiratory system, urinary system, and gastrointestinal tract are usually evaluated secondarily or through necropsy. Secondary effects include shallow breathing, regurgitation, agitation on urination, and eye rubbing, which are broadly categorized as reflex behavior, and blood in feces, urine, and sputum, or nasal discharge, which are more indicative of the primary irritant effect. These results are not quantifiable by present methodologies. Primary evaluation through necropsy is also a qualitative procedure and has not undergone the extent of standardization and validation that the skin and eye tests have. Other methods for testing irritation such as resistance/compliance tests of the pulmonary system, direct observation by scope of the esophagus

and gastrointestinal tract, and roentgenographic examination with and without radio-opaque dyes, have not received wide use as testing techniques for Government regulatory purposes.

Sensitization

Some substances, although not necessarily primary irritants, elicit an irritant-type response after repeated contact with the organism. Tests for this sensitization potential involve exposing the animal to an agent at doses below those necessary to produce signs of primary irritation, waiting an appropriate interval, and then challenging the animal with the substance again at a different site (11). If the response on retest is substantially higher than the initial response, the agent can be classified as a sensitizer. Various test methods and modifications have been proposed for testing sensitizing potential (11, 13-21). Perhaps the most common direct test for sensitization is the guinea pig maximization test. In this test, the agent is presented in Freund's complete adjuvant which increases the response. Studies on the mechanism of sensitization have shown that the agent or a metabolize of it (antigen) may induce the lymphocytes of the body to form a complex molecule (antibody) which reacts with the antigen to form an antigen-antibody complex. This reaction may be with circulating free antibody or with lymphocyte-bound antibody, The formation of the antigen-antibody complex induces the production and release of histamine and other compounds which cause the erythema and edema at the site of antigen attack, or may cause anaphylaxis if the antigen reaches the blood stream (22,23). The problem with testing methods based on this mechanism, such as immunoelectrophoresis, radioimmunoassay, ring test, hemagglutination tests, or microphage migration, is that they do not measure the actual adverse effect [dermatitis or shock) but measure an indicator response. The methods are valuable, however, in demonstrating the presence of antibody capable of producing these health effects. Several comparative tests have demonstrated that these in vitro techniques are often more sensitive indicators of the hazard than the classic in vivo ones (22,23).

Structural Effects

The basic determination of structural effects on organs and systems begins with the determination of absolute and relative weight changes. Decreases in absolute body weight or rates of weight gain for a test with a substance incorporated into the food or water may show either that the agent is unpalatable or that it is interfering with the energy balance, the central nervous system (CNS) regulation of food or water consumption, or the motivation of the animal. Although substances administered by other routes of exposure may also interfere with palatability of food, through direct or indirect effects on the sensory nerves, this is less common. Increases in weight may be caused by a proliferating tumor mass. The evaluations of structural changes can be obtained from animals exposed during subacute experiments.

In general, subchronic or subacute experiments are designed to last approximately 10 percent of the animals' lifespan (90 days for rats). Immediately preceding and during the experimental period, observations on animals should include rate of growth, food and water consumption, demeanor, and reflex behavior; blood, urine, and feces should be collected. During the experimental period, tissue biopsies may be taken for observation of structural changes. These techniques may be unreliable, however, if a structural change is localized and not included in the biopsy material, and such manipulation is often not allowed by regulatory testing guidelines. If biopsies are done, additional animals are required to maintain the statistical validity of the experiment. At the end of the experimental period, the animals are sacrificed and the organs are inspected for gross changes, removed and weighed, and preserved for histologic treatment and microscopic examination (5). The specific organs and tissues removed and examined will depend somewhat on the expected action of the agent administered (usually perceived from preliminary testing) but should include at least the brain, liver, kidneys, spleen, heart, testes (and epididymis) or ovaries (and uterus), thyroid, and adrenals.

Changes in organ weights may signal a functional change in this organ or in other organs; for example, an increase in heart weight could be due to a decrease in oxygen diffusion from the lungs, an increase in adrenal weight could signal a blockage of steroid synthesis within it, etc. The weights of organs can be directly compared with those from control animals; however, this often introduces an artifact since experimental and control body weights are usually different. It is common practice therefore to determine the relative weights of the organs in relation to the total body weight of the animal. Recently it has been proposed that the relative weights should be taken as a function of the animal's brain weight, the postulation being that the brain's growth curve devi-

ates the least of any tissue in the body. While normalization based on this procedure would tend to emphasize changes more than other techniques currently in use, it is not yet widely accepted. Tissue dry weight, after desiccation or ashing, has also been used as a tool for determining mechanisms of growth and metabolic balances (24). This method has a major drawback, however, because it removes the organ from further studies such as microscopic examination.

After gross observation and weighing, the organs and tissues are preserved for histological preparation and microscopic examination. The most common methods for the preparation and staining of individual tissues involve fixing with 10-percent buffered formalin solution, embedding in parafin, and staining with hemotoxylin/eosin, Many pathologists prefer other fixing and embedding media, and certain tissues require different procedures. There are also special stains for highlighting different cellular components. There is no one best method for preparation and observation of the tissues. The most valuable procedure from the pathologists' viewpoint is to prepare the tissues in a number of ways, which allows comparison of various aspects such as specific cellular components, nuclei, cell membranes, etc. (25-27).

Special consideration can also be given to techniques in histochemistry and electron microscopy. These methods are not used routinely in toxicological evaluation and depend on a knowledge of the mode of action of the toxic agent. They can, however, indicate changes in cellular metabolism or structure before those changes become manifest by the conventional histological procedures, and therefore they may be more suitable for observing changes from agents whose toxicities are low or develop slowly. The equipment necessary for these techniques is generally more expensive than that needed for the more conventional microtechnique methods. They are also more time consuming and less standardized than conventional methods. The histochemical methods, although they might be more appropriately classified as tests of organ function, are becoming more widely accepted with investigators studying mechanisms of toxic: action.

Functional Effects

Frequently, changes in organ or system function are observable before any change in structure becomes apparent. The test methods discussed below have generally been adapted from human to animal use, and results are ordinarily compared with animal control values and are not necessarily comparable between species, Methods for evaluation of pulmonary function (28), cardiovascular function (29), and brain and neural activity (30) have been modified for human and animal use. These methods include testing ventilator flow, resistance, compliance, and gas diffusion capacity for the pulmonary system; electrical activity of the heart, and blood flow and pressure for the cardiovascular system; electrical activity of the brain (field and single unit) and muscles; perception threshold, reflexes, and chronaxy for neural function. The significance of changes in brain activity as a determinant of toxicity is under question at present, however.

Generally, in these evaluations, each animal serves as its own control. Baseline data for each procedure is determined prior to administration of the agent, and any changes in function are noted during and after administration, since it is important to determine whether the agent causes reversible or irreversible changes in function.

Various other methods of testing for organ or system function rely on both primary and secondary parameters. For example, liver function may be assessed by dye clearance studies (primary) or by analysis of serum enzyme concentrations (secondary). Most of the secondary procedures are now automated and available through various clinical laboratories at a reasonable cost, Many investigators, however, still prefer to perform the tests manually, and standard procedures are well-defined and available in several texts (31-33). Various modifications of these tests for specific animal systems have been developed and published, Tests usually considered appropriate include total and differential blood counts, serum enzyme and ion analysis, urinalysis (especially for metabolizes of the agent), and liver and kidney function tests (dye clearance). The value of these tests is that abnormal results will often precede obvious structural damage of the organ system in question and will be apparent at lower dose levels.

The study of hematologic effects encompasses changes in the bone marrow as well as those in the cells of the circulating blood. Observations are made of the cells and of their absolute and relative numbers. Specific tests such as dye dilution for blood volume, specific gravity, sedimentation rates, osmotic fragility, hematocrits, or clotting time, are not routinely performed but may be indicated. Serial bone marrow biopsies may also be performed for hematologic effects; the results

give both structural and functional information, but these techniques are also not widely used.

The functions of the liver may be tested for biliary obstruction (icterus index, alkaline phosphatase), liver damage (thymol turbidity, plasma protein ratios, cholesterol ratio, glucose level, transaminase level, and cholinesterase level), excretory function (bromsulphalein clearance, bilirubin tolerance), and metabolic function (glucose tolerance, galactose clearance), The most common tests used in toxicology are the serum alkaline phosphatase and serum transaminases, and in some cases a dye clearance (bromsulphalein) or glucose tolerance.

The kidneys are responsible for excretion of certain substances, e.g., urea, and for concentration and dilution of urine. Tests for excretion involve dyes like phenolsulphonphthalein and also measure such substances as urea and creatinine, Concentration and dilution tests involve measurement of urine specific gravity after fasting for various periods. These tests are generally not used in toxicology screening studies unless there is reason to believe the toxicant acts on the kidneys.

The evaluation of these tests may proceed with or without modification of the tissue metabolism. That is, promotors and inhibitors of enzyme systems, e.g., SKF-525A for mixed function oxidase, may be used to enhance the susceptibility of a particular organ or system to damage from a toxicant. This in effect maximizes the response so that the toxic action can be more readily observed.

Many other specific tests are available for evaluating various organs and systems such as sperm motility, specific gravity of cerebrospinal fluid. calcium-phosphorus ratios for the skeletal system along with tensile strength and compaction, epinephrine sensitivity of heart muscle, acetylcholine test of lungs, metabolism of excised tissue, work and strain measurements of the various muscle systems, etc.

Behavioral Effects

A recent addition to the field of toxicology has been behavioral testing. Testing methods have been devised for everything from simple perception to complex tasks involving perception, learning, judgment, motivation, and motor activity. The value of the behavioral methods lies in the ability of the nervous system to respond to toxic agents a t doses much lower than those necessary to produce "classic" signs of toxicity in the organism. Therefore, these methods are a potential sensitive indicator of hazard and can be used as an "early warning system. " Several good reviews of behavioral toxicology are available (34-37).

Some of these methods rely on newer methods of analysis such as contingent negative variation (CNV). Some are really an application of preexisting principles such as dorsal evoked potentials and neuromuscular transmission time that have been widely used by experimenters in the field of neurophysiology. Most of these techniques have only recently been turned to the evaluation of toxicity.

At the present time, standardization and validation of behavioral techniques has not been accomplished. The question often raised by regulatory agencies is how do you relate an observed behavioral decrement to an adverse health effect, especially if there is no concurrent structural change apparent in the nervous system. Because of these factors, behavioral studies are often derogated by these agencies when setting exposure limits for toxicants. Current research is being conducted, however, under Government contracts to answer some of these questions.

Comparison of Short- and Long-Term Methods for Systemic Toxicity

Most of the procedures noted in this section are equally applicable to short- and long-term testing, obvious exceptions being irritation and sensitization tests. The value of long-term testing for systemic toxicity lies in the ability to use low doses that do not produce detectable adverse effects in a short time period to see whether bioaccumulation and cumulative effects occur. Predictions of the effect of bioaccumulation can be made knowing the effects of short-term high doses, but final evaluation of the toxicity depends on the long-term effects observed, As will be pointed out in the metabolism section, toxicants can be potentiated or inhibited by the metabolism and relative accumulation of the toxic moiety. Without definitely knowing the various metabolic reactions, rates, and probabilities, it is impossible to accurately predict toxic effects. High short-term doses may induce a toxic reaction, such as death from pulmonary edema, that might mask long-term, low-level exposure effects such as liver cancer, The differences in short-term and long-term tests for systemic toxicity include the number of interim measurements allowed, the ability to ascertain the types of effects which might develop only over a long period and the progression or time course of toxic manifestations, and

the ability to evaluate mechanisms of bioaccumulation or adaptation in the organism.

Mutagenicity

Rapid identification of a food contaminant as a possible mutagen is necessary to reduce the potential genetic risk to humans who might contact the contaminant. Mutagenic effects on humans often cannot be directly detected, and deleterious effects on the human gene pool may not become apparent for many generations if, for instance, the deleterious effect is due to a recessive gene. Heritable genetic damage in humans may result from any of several types of effects on the genetic material, The two major classes of effects are point mutations, which generally affect a single gene or part of a gene, and more extensive chromosomal effects such as gross changes in structure or changes in number.

Only a few tests are available that directly evaluate genetic effects of exposure of mammals to chemicals: however, the potential of a chemical to produce heritable genetic alterations in man can be evaluated indirectly from its effects on genetic material in various biologic test systems, including micro-organisms, mammalian cell cultures, insects, and intact mammals.

For substances that cannot feasibly be eliminated from the human environment, it is not sufficient to identify the existence of a genetic hazard; quantitative assessment of the risk involved is necessary for appropriate regulatory activity, such as establishing action levels or tolerances for food contaminants.

Mutagenicity testing is also used to prescreen chemicals as an indicator of carcinogenic potential and, less frequently, other toxic effects such as teratogenicity. This application is based on empirical demonstration or correlation between mutagenicity and carcinogenicity of chemicals (38) and does not depend on the assumption that the same mechanism is involved in both types of effect.

Approaches to Testing

As in all toxicological tests, mutagenicity tests may produce false negatives (a negative result when the substance is actually mutagenic) and false positives (a positive result when it is not mutagenic), and correlation between the results from two test systems may be poor. Ideally, a mutagenicity test system should be sensitive enough to detect any chemical that may cause heritable genetic damage and its results should be reproducible. Finally, the test results should be quantitatively applicable to mutagenesis in humans. Since no single test can fulfill these requirements and none is reliable enough to stand alone as an indicator of mutagenic potential, mutagenicity ‘testing should include a variety of systems selected to show whether the test substance or its metabolize produce any of a range of genetic effects. The test battery approach includes systems that will detect several types of gene mutations, chromosomal aberrations, and DNA repair; thus, this approach offers the greatest reliability for determining mutagenic potential, Tests that evaluate effects in intact mammals are essential for predicting mutagenicity in humans.

Since screening large numbers of chemicals by the test battery approach (39,40) may be prohibitively costly, a hierarchical approach to mutagenicity testing, known as tier testing (41), has been suggested. Tier 1 consists of relatively inexpensive short-term prescreening tests. These use micro-organisms or other in vitro systems to determine priorities for indepth testing. Substances that produce positive results in these tests, as well as those that are negative but are structurally similar to known mutagens or to which there is a substantial risk of exposure for humans during or preceding their reproductive years, should continue into Tier 2.

Tier 2 tests are usually designed to detect substances that are not mutagenic in vitro but are metabolized to an active form in the intact mammal. Tests used at this level may include the dominant lethal test, in vivo cytogenetic tests, the host-mediated assay, and body-fluid analysis. Substances that are negative in Tiers 1 and 2 are generally considered safe for use and are given very low priority for further testing.

Only substances for which it is important to assess risk are subjected to Tier 3 testing, designed to permit quantitative evaluation of mutagenic potential. Tests used at this level include multigeneration mammalian studies, such as the heritable translocation test, X chromosome loss test, and specific loci test in mice.

Tier testing may represent an efficient use of resources in large-scale mutagenicity testing, but the use of prescreening tests carries serious disadvantages in determining mutagenic potential, If test chemicals are prescreened by a single microbial test, the proportion of false negatives may be unacceptably high, and potentially hazardous or useful substances may escape further testing. The use of two or three tests at this level, including both micro-organisms and mammalian

cell cultures with and without activation by mammalian enzyme systems, may substantially increase the reliability of prescreening (96). Nevertheless, such cell systems may not approximate metabolic events in the intact mammal closely enough to reveal the mutagenic action of some substances that are potential human mutagens.

Whatever the testing approach and test systems selected, mutagenicity tests should include a positive control as well as negative (untreated and solvent) controls. The positive control substance, a known mutagen in animal systems that is selected for its structural similarity to the test chemical, serves to demonstrate the sensitivity of the test organism and the efficacy of the metabolic activation system used.

Current Test Systems

Chromosomal effects of many substances have been demonstrated in plants such as *Vicia faba* and *Tradescantia* (42), and the latter organism has also been used in detection of somatic mutation (43). While the genetic events involved (alterations in DNA) are the same as those in mammalian cells, their relevance to human mutagenesis has been questioned because of the major phylogenetic and physiologic differences between plants and animals; thus, a negative result in plants does not indicate that a substance is not a mutagen in mammalian systems.

Of the many bacterial species that have been used to detect point mutations, the most extensively employed are the Salmonella typhimurium mutants developed by Ames (44,45). The Ames test uses a series of histidine-requiring mutant strains that revert to histidine-independence by specific mechanisms, either base-pair substitutions or frameshift mutations, The original strains have undergone several further modifications that increase their sensitivity to mutagens by interfering with DNA repair or modifying the cell wall to enhance the penetration of chemicals into the cell. Bacteria treated with the test chemical are plated on selective media or cultured in liquid suspension to determine the number of revertants. A reproducible mutation rate twice the spontaneous (control) **rate** is usually considered evidence of mutagenic activity.

Because microbial cell systems do not possess the metabolic capabilities of mammals, they will not detect chemicals that exert a mutagenic effect through metabolic intermediates, Several activating systems have been developed for use with in vitro test systems to duplicate the effects of mammalian metabolism. The most extensively used means of metabolic activation is the addition of microsomal mixed-function oxidase enzymes, typically from rodent liver homogenates, to metabolize the test chemical in vitro, This activating system is added to the culture medium as part of the Ames testing procedure with S. *typhimurium*, and it has provided evidence for the mutagenicity of many substances that have no direct mutagenic effect on these bacteria **(51).** Microsomal enzyme activation is also used with other microbial test systems (46-48). The major drawbacks of this system are that it would not detect chemicals metabolized to mutagenic intermediates by mechanisms other than liver microsomal enzymes, e.g., substances metabolized by the intestinal flora, and it is possible that the in vitro metabolism of the substance does not adequately mimic its metabolism in the intact organism because of competing reactions. Another drawback is that a standardized in vitro activation system has not been devised to date.

Another widely used bacterial system is the multipurpose strain of *Escherichia coli* developed by Mohn and coworkers (49). This strain can be used to measure reverse mutations restoring the ability of the bacteria to synthesize the nutrients arginine and niacin. Forward mutation rates in two genes controlling galactose metabolism can also be scored in this strain of E. coli. Use of this test organism permits the detection of several types of mutation in a single experiment.

Eukaryotic micro-organisms that are used to detect the ability of chemicals to produce point mutations include haploid strains of the yeasts Saccharomyces cerevisiae and Schizosaccharo-*myces* pombe (47,50) and of the ascomycete *Neu*-rospora crassa (51). A diploid strain of S. cerevisiae permits detection of chromosomal damage expressed as mitotic recombination that produces phenotypic color changes (52).

Whole-animal activation mechanisms can circumvent this problem but they are generally much less sensitive than tests using in vitro activation. In these systems, rodents are exposed to the test chemical by an appropriate route, and the effect of rodent metabolizes on microbial genetic markers is determined by either body-fluid analysis or host-mediated assay procedures,

In the body-fluid analysis (53-56), the microorganisms are treated with the urine, blood, or homogenized tissues of the exposed animals. Caution is necessary in interpreting negative results of these studies, in the absence of supplementary pharmacologic data, since even if the chemical is metabolized to a mutagen, other factors such as

tissue-specific activation and detoxification mechanisms and the half-life of the compound and its metabolizes may affect test results. In the host-mediated assay (57,58), the micro-organisms are exposed to mammalian metabolic products of the test substance by being introduced into the peritoneal cavity, circulatory system, or testes of the host mammal. The host is treated with the test substance, and after an appropriate incubation period, the indicator organism is removed and examined for mutations.

Genetic damage in micro-organisms can also be assessed indirectly through the use of DNA repair-deficient strains of bacteria (44,59). These tests organisms and otherwise identical strains that have normal ability to repair DNA are treated with the test substance. Toxic action of the test substance produces zones in which bacterial growth is inhibited, and the difference in size between the inhibition zones in repair-deficient and normal strains indicates the extent to which this toxicity is due to damage to the DNA.

A number of mammalian cell systems in culture have been developed for detecting point mutations, including cell lines derived from mouse lymphomas, Chinese hamster ovaries and embryos, and human fibroblasts and lymphoblasts (60-63). In addition, gross chromosomal changes such as breaks, gaps, and rearrangements can be microscopically observed in these cells. Stable rearrangements, such as translocations and inversions, are considered evidence of heritable changes. The induction of only gaps and breaks is not regarded as evidence of mutagenicity, because these aberrations often occur as a result of general cytotoxicity and thus may be present only in moribund cells. Like microbial systems, in vitro mammalian cell tests can be used in conjunction with activation by mammalian enzymes or with whole-animal activation to permit detection of mutagenic effects by metabolic products of a test substance.

Mammalian cells, including human white blood cells, are also used to detect chemical damage to DNA by measuring unscheduled DNA synthesis (63). This indirect indicator of genetic damage is evaluated by measuring the uptake of radioactive thymidine for repair of damaged DNA during those stages of cell growth when DNA synthesis does not normally occur. A similar test with mouse spermatocytes exposed in vitro or in vivo demonstrates effects on DNA in germinal cells (64),

Sister chromatid exchange, a reciprocal exchange of segments at homologous loci. measured by autoradiographic methods, has also been used to examine a variety of chemicals (65,66). Sister chromatid exchange in various cell systems has been demonstrated following exposure to known mutagens, but additional work is needed to define the extent to which results are correlated with more traditional mutagenicity tests,

The fruit fly Drosophila melanogaster is used in a comprehensive and extensively characterized mutagenicity test system which can detect all types of mutagenic activity at a fraction of the time and cost of in vivo mammalian testing (67). The large number of genetic markers and known chromosomal aberrations make it possible to assay a chemical for many types of mutagenic activity in a single test. The sex-linked recessive lethal test (68) in Drosophila is a very efficient mutagenicity assay, since about 20 percent of the insect's genetic material is located in the X chromosome. Recessive lethal changes caused by point or chromosomal mutation can be mapped and in most cases the nature of the change causing the mutation can be determined. This test also permits a quantitative assessment of mutagenic activity. Because of the large size of Drosophila chromosomes, this organism can also be used readily to assess meiotic and mitotic recombination, dominant lethality, translocations, and deletions. Some indirect mutagens that required metabolic activation have been shown to be mutagenic in Drosophila, indicating that these insects have a microsomal mixed-function oxidase system (69). However, to determine whether Drosophila test results are useful for risk assessment, more information is needed on how their metabolism of foreign chemicals compares to that in humans. A few other species of insects have also proven useful in mutagenicity testing, including 'several species of the parasitic wasp, Habrobracon (70).

Clearly the most accurate predictions of mutagenic potential in humans can be drawn from tests that determine direct genotypic and phenotypic effects in mammals exposed to the test chemical by routes relevant to human exposures, Several direct mammalian tests exist, but these have the disadvantage of detecting only a few of the possible types of genetic damage. Chromosomal damage occurring in vivo can be detected in several different cell types, such as bone marrow cells and circulating lymphocytes, and the presence of micronuclei in red blood cells (7 I). Cytogenetic tests using mammalian lymphocytes and the micronucleus test offer the advantage of permitting direct comparison with effects on humans resulting from accidental exposures; how-

ever, these tests demonstrate only effects on somatic cells and do not provide direct evidence of heritability.

Cytogenetic changes in mammals can also be evaluated in germinal tissue from the testes. In the direct spermatocyte test (72), male mice are exposed to the test substance. After sufficient time for the treated spermatogonia to reach the spermatocyte stage, they are examined for cytogenetic abnormalities. This test allows for the actual observation of induced cytologic changes in premeiotic male germ cells, but it does not permit detection of effects in postmeiotic cells or their transmission to the offspring.

Effects on offspring can be evaluated in mice by the heritable translocation test (73) and the X chromosome loss test (74). In the heritable translocation test, F_1 male offspring of treated mice are mated to determine sterility, and indication o possible translocation heterozygosity. Chromosomal effects are then confirmed by cytogenetic analysis of the germinal cells of the male of offspring. The X chromosome loss test permits the detection of chromosome loss resulting from nondisjunction in the female, since, unlike somatic chromosome aneuploids, animals of XO genotype are usually viable, Aneuploidy for the X chromosome can be detected by genetic markers and confirmed by cytologic observations.

The dominant lethal assay, usually performed in the rat or mouse, uses fetal loss as an indicator of induced chromosomal mutations in male germinal cells (75). The death of the zygote is assumed to result from chromosomal abnormalities in the sperm of male mice exposed to the test chemical. This test is relatively easy to perform and its results have been positively correlated with mutagenicity in other animal systems. Preimplantation loss alone is not used as an indication of mutagenicity since it has been found to occur for reasons other than chromosomal changes in the sperm. Disadvantages of this test are its relative insensitivity and difficulty in clearly distinguishing weakly positive results.

Only one test is available at present that can detect heritable gene mutations induced in mammalian germ cells. In the specific locus assay in mice, forward mutations at seven loci, affecting characteristics such as coat and eye color are mated with mice homozygous for recessive alleles at these loci (75). Because such a small number of loci are involved, this test required the scoring of 20,000 to 30,000 offspring at each dose level to produce reliable results and is therefore very costly and time consuming.

Carcinogenicity

In the event of massive or long-term environmental contamination of food destined for human consumption, one of the decisions to be made is whether the contaminant appears to pose a significant carcinogenic risk. With this in mind FDA submits the candidate compound to the Chemical Selection Working Group at the National Cancer Institute for consideration under the carcinogen bioassay screening program (76,77).

Prerequisites for a Carcinogenicity Study

Once the compound has been selected, it is screened using a chronic or lifetime exposure regimen (76,78,79). However, before the long-term study is undertaken, specific toxicologic profiles must be obtained. Young healthy adult animals of each sex and strain to be used in the long-term studies should be used in the preliminary studies. The animals should be of uniform age and weight and should be tested using the same formulation and route of exposure to be used in the long-term studies. The first is an acute study designed to gain additional information on the acute toxicity, assuming there is a paucity of data on this aspect of toxicity, and to determine the lethality of the test compound. The duration of this test should not exceed 24 hours and should include at least three dose levels determined by a geometric progression. one of the dose levels selected should represent the highest dose to be used in subsequent studies. Throughout the investigation, all relevant clinical signs should be recorded. Necropsies should be performed on a random selection of animals of each sex and strain, and any abnormal histopathologic changes should be noted.

After the 24-hour study, a 14-day investigation should be initiated in an effort to ascertain the doses necessary for the subchronic study, the next prerequisite investigation for the chronic study. This toxicologic study requires five dose levels, with the highest one, estimated from the 24-hour acute study, producing no more than 10-percent lethality. The other dose levels should represent geometric decrements of the highest dose. Animals should be treated with the test substance for no more than 14 days, held another 24 hours, and then sacrificed for necropsy. Throughout the study, the animals should be observed for clinical signs of toxicity. Other toxicity data, such as those derived from organ function tests and metabolism studies, are also necessary.

The next toxicity study involves the administration of the test substance for 90 days and is used as a predictor of the maximum tolerated dose (MTD). This can be defined as the highest dose given during a chronic study that can be predicted to not alter the animals' normal longevity from effects other than carcinogenicity, In practice, MTD is considered to be the highest dose that causes no more than a 10-percent decrement in weight compared to controls. Five dose levels are required in this study, with a minimum of 10 animals of each sex and strain in each dose group. The highest dose level used should be the lowest concentration that produced any detectable untoward toxic effects in the 14-day study. The remaining dose levels should be determined as in the 14-day study, If the selected dose levels do not produce a discernible no-effect level, the study should be repeated with lower doses.

Carcinogenic Bioassay

The chronic study (76,78,79) represents the essence of the carcinogenicity bioassay. It is used to determine the carcinogenicity of a compound in males and females of two mammalian species, usually the rat and mouse. The species selected are tested throughout their entire lifespan. Each test group should consist of a statistically representative number of animals. The highest selected dose should represent MTD and the remaining dose levels should be adjusted accordingly. There should be at least one control group, in which the animals receive only the vehicle used for administration of the test material, If no vehicle is used, this control group should be untreated but identical in every way to the experimental groups. In addition to the concurrent control group, a colony or historical group should be used for the comparison of longevity, spontaneous diseases, and spontaneous tumor incidence. The historical control may also be used for statistical comparisons. In some studies, a positive control group that has been treated with a compound structurally similar to the test compound and known to be carcinogenic in the test species may be indicated. However, because of the added risk of handling a known carcinogen, a positive control group is seldom used.

Throughout the study, animals must be observed for signs of toxicity. Every animal should be examined carefully each week, Animals should be weighed and food consumption measured. In some cases, it is desirable to evaluate tissue distribution and concentration of the substance or its metabolizes.

The animals in any one test group should be sacrificed at an adjusted or prearranged date. However, a group can be terminated earlier if there has been high cumulative mortality. Moribund animals should be sacrificed immediately upon discovery to lessen the likelihood of unobserved deaths and subsequent autolysis or cannibalism. Control groups should be sacrificed according to the original or adjusted sacrifice date; the later date is preferred. Because of the strong dependency on histopathologic results and to avoid possible criticisms of the study, necropsies and histopathologic examinations should follow standard procedures required by regulatory guidelines.

The major drawback to the carcinogenic bioassay procedure is the time and cost required to complete and analyze such a study. If the study is done properly, however, the results should be conclusive and for the most part indisputable, although there still remains the question of extrapolation of results to the human population.

Short-Term Testing as a Prediction of Carcinogenicity

Evaluation of carcinogenicity has generally relied on the results of long-term animal studies, To use this kind of testing approach for every substance suspected of being carcinogenic would be cost-prohibitive and certainly impractical in terms of the overall time required to test all suspected chemicals. Therefore, short-term tests are being developed to identify carcinogenic substances. There has been much criticism concerning the comparison of short-term testing results from different laboratories because of the varying conditions and refinements in techniques practiced among testing facilities. The protocols for these short-term tests, especially those involving mammalian enzyme activation systems, have not been standardized or validated through interlaboratory comparative testing procedures; thus, comparisons of the data from one laboratory to the next have often produced conflicting conclusions. In this light, a study conducted in one laboratory that compares several short-term testing systems has added importance in clarifying the relative usefulness of the compared systems,

In a recent study by Dr. Ian Purchase (80), 120 organic chemicals (50 known carcinogens and 62 noncarcinogens, based on published experimental data) were evaluated for activity in six short-term test systems. These systems included: 1) mutation of Salmonella typhimurium (45), 2) cell transformation (81), 3) degranulation of endoplas-

mic reticulum (82), 4) sebaceous gland suppression (83), 5) tetrazolium reduction (84), and 6) lesion formation after subcutaneous implant (85). Four additional tests used in a preliminary study were found to be insufficiently accurate or sensitive to justify a full evaluation. The tests rejected were transplacental blastomagenesis (86), piperidine alkylation (87], iodine test (88), and the acridine test (89).

Although there were considerable variations between tests in their ability to predict carcinogenicity, two tests were quite accurate in distinguishing between the known carcinogens and the noncarcinogens. These were the cell transformation test and the bacterial mutation test, which had accuracies of 94 and 93 percent, respectively. The use of cell transformation and bacterial mutation together provided an advantage over the use of either alone, predicting 99.19 percent of carcinogens. Not surprisingly, the inclusion of the other four tests in a screening battery with these two resulted in an improved ability to detect carcinogens (99.97 percent), but greatly decreased the accuracy and discriminatory value of the battery. It is important to note that all tests generated both false positives and false negatives: the percentages for both can be readily calculated by subtracting the positive predictability values from 100 percent.

A description of each of the tests, with comparative percent accuracies for predicting carcinogenicity as determined by Purchase et al. (80) is as follows:

- Bacterial mutation. The procedures used were those of Ames, in which four strains of S. typhimurium (TA 1535, TA 1538, TA 98, TA 100) were tested with each compound in an assay medium containing a metabolic activation system composed of rat liver postmitochondrial supernatant (S-9 fraction) and cofactors. The overall accuracy of the test in predicting the carcinogenicity of the compounds in this study (80) was 91 percent for carcinogens and 94 percent for noncarcinogens. These figures agree with the previously published value of 90 percent by McCann et al. (90) but are considerably higher than those published by Heddle and Bruce (91) who found 65 and 81 percent, respectively.
- Cell transformation. The procedures used were those of Styles (81) and involved three types of mammalian cells—human diploid lung fibroblasts (WI-38), human liver-derived cells (Chang), and baby Syrian hamster kidney cells (BHK 21/cl 13). In all assays, the cells were used with and without metabolic activation with S-9 fraction as described previously. Without activation, very few carcinogens transformed hamster or human cells in the period of study. With activation, all cell lines detected carcinogens with an accuracy of 88 percent or better. Furthermore, by the use of both the hamster cells and either of the human cells, the overall accuracy was improved to 94 percent (91 percent for carcinogens and 97 percent for noncarcinogens).
- Degranulation. The procedure used is that of Williams and Rabin (82) and the test is commonly called the Rabin test. The test measures the loss of ribosomes (degranulation] from isolated rat liver endoplasmic reticulum following incubation with the test compound. The overall predictive value was 71 percent for both carcinogens and noncarcinogens.
- Sebaceous gland test. The procedures used were those of Bock and Mund (83) where test chemicals were applied directly to the skin of mice and a depression in the ratio of sebaceous gland to hair follicles indicates a positive response. The overall predictive value of the test was 65 percent (67 percent for carcinogens and 64 percent for noncarcinogens).
- Tetrazolium reduction. The procedures used were based on those described by Iversen and Evensen (84). Test solutions were applied directly to the skin of mice and the skin samples were incubated in tetrazolium red solution. An increase in the in situ biologic reduction of the colorless tetrazolium to a colored formazan compound indicates a positive response. The overall predictive value of this test was 57 percent (40 percent for carcinogens and 71 percent for noncarcinogens).
- Subcutaneous implant. The procedures used were essentially those of Longstaff and Westwood (85] and involved the subcutaneous implantation of a filter disc overlaid with a gelatinous suspension of the test compounds into mice. After 3 months, the surrounding tissues were scored for lesions. The overall predictive value for this test was 68 percent (37 percent for carcinogens and 95 percent for noncarcinogens).

While the work of Purchase and associates (80) does present a very salient assessment of the more pertinent in vitro assays, it also points out

the shortcomings of this type of approach. Short-term tests do not use the induction of cancer as an endpoint, but each has a parameter, such as induction of a point mutation, that varies with the carcinogenicity or noncarcinogenicity of the test substance. Accordingly, the authors feel that these test parameters should not be given a greater weight than that of any other arbitrary response, regardless of how biologically significant any of these tests might appear to be with respect to the theories of the chemical induction of cancer. Also, however much generalized data might be generated to support the predictive accuracy of the given test, this accuracy should not be assumed to apply uniformly to compounds of every chemical class,

Teratology and Effects on Reproduction

An investigation of a teratogenic agent involves the study of congenital malformations other than those that are inherited. Teratogens themselves act as triggers for malformation induction. Malformations may include gross, histological, molecular, and behavioral anomalies,

The sensitivity of an animal to a teratogen is determined by: 1) the period in which the insult is received during the gestation period (this includes before germ layer formation and during embryogenesis or organogenesis); 2) the dose and the route of administration of the compound; 3) placental transfer of the suspect teratogen, including its lipid volubility, protein binding ability, and metabolism, and 4) uterine and dietary factors.

A toxicologic profile of a suspect teratogen requires an evaluation of potential hazards to reproduction and, particularly, to developmental processes that respond to environmental insult through mutation, chromosomal aberrations, mitotic interference, altered nucleic acid synthesis, enzyme inhibition, and altered membrane characteristics. This is particularly important in the conceptus, embryo, and the neonate where the biochemical, morphologic, and physiologic properties change rapidly. Thus risk assessment must not only address the teratogenicity of a contaminant but also its effect on reproduction.

Classical Approach

For an adequate teratologic assessment (79,92, 93), exposure to the toxicant should parallel as closely as possible that expected in the human population. The pharmacologic activity of the compound as well as its acute and chronic toxicity is also a consideration. For teratogenic studies, the toxicant is usually administered daily on the specific days of gestation representing the period of greatest sensitivity. Administration of the test substance should begin at or before implantation and should continue throughout the period of major organogenesis.

The selection of an animal species for evaluation of teratogenicity is an important consideration. Test protocols currently in use recommend at least two mammalian species, the first being a rodent (e. g., mouse. rat, hamster) and a nonrodent mammalian species (e.g., rabbit). One species should be the same as that used in the test for reproductive effects.

In conducting the teratogenic investigation, at least three dose levels should be used, along with concurrent control groups. These control groups should consist of untreated animals, animals treated with the vehicle of administration only, and, possibly, animals treated with an agent that is known to cause the effect that is being investigated. The use of historical or colony controls may also be helpful in evaluating the data. Both test and control animals must be young, mature, prima gravida females of uniform age, size, and parity. The control groups should be handled and maintained like the test groups. The highest dose level to be considered should produce signs of embryo or fetotoxicity as suggested by fetal growth retardation and more significantly by maternal or fetal mortality; however, maternal mortality should not exceed about 10 percent. The other dose levels can be obtained in a decreasing logarithmic fashion to a suspected no-observable-adverse-effect level. In the classical teratology study, treatment may be by gavage so that the administered dose is accurately known. It has been shown, however, that use of this route of administration may induce anomalies in progeny that are a consequence not of the test compound but of the stress of dose administration (94).

Prior to initiating the study, it is important to determine whether sires and dams have successfully mated, This includes examinations for the presence of plugs or evidence of sperm in vaginal smears. Throughout the study, females should be observed for behavioral changes, food and water consumption, body weight, vaginal bleeding indicating possible abortion, and spontaneous deaths. Females showing signs of aborting or delivering prematurely should be sacrificed.

Fetuses should be obtained 24 hours before anticipated parturition by cesarian section, Dams should be sacrificed and a complete necropsy per-

formed with emphasis on the reproductive organs. Data to be obtained include maternal weight and weight of the gravid uterus, the number of fetuses, location of fetuses within the uterine horn, number of corpora lutea, the number of sites of resorptions (either early or late), the number and weights of live and dead or moribund fetuses, the number of dams having dead fetuses or showing reproductive wastage, and any internal or external anomalies through appropriate procedures (complete necropsy or Alizarin Red staining). Approximately one-half to two-thirds of each litter should be screened for skeletal anomalies while the remainder of the litter should be screened for aberrant soft-tissue effects. The significance of each observed anomaly should be evaluated relative to its natural incidence and its occurrence in controls.

Effects on Reproduction

The intent of the reproductive study is to determine whether the substance under consideration produces demonstrable effects from mating through lactation, in pregnancy, or in growth and development of progeny from conception through maturity (79,92).

Reproductive studies are normally performed in at least two mammalian species, one of which should be the same as one of the two species used for teratogenic evaluation. For rodents, at least three dose levels are selected in addition to the controls. The highest dose must produce an observable effect but no more than 10-percent lethality, whereas the lowest dose should yield no demonstrable adverse effects. The control groups used in reproduction studies are similar to those discussed for other studies, namely a vehicle control and possibly a historical control.

The period of dosing and observation for these studies is rather long, i.e., about 1.5 years. The test material is administered to the F_0 and the F_1 generations. The third generation of animals (F_2) is then exposed *in utero* and through weaning.

Dosing of the F_0 generation should begin as soon as possible after weaning and definitely before the animals (males and females) are approximately 6 weeks of age. The substance under test should be administered daily to the F_0 generation for approximately 100 days before breeding; dosing should continue in the females that have been bred and have delivered progeny until all of the F_1 generation have been weaned.

At approximately 30 days of age, a statistically valid number of animals, males and females, from the F_1 generation are selected and exposed to the test material for at least 120 days. At this point they are bred to produce the F_2 generation.

The types of data to be collected include growth and time of delivery for each weanling as well as overt signs of toxicity, the general behavior and condition of the mothers, measurements of spermatogenesis in all F_1 generation males used to produce the F_2 generation, litter size, number of stillborn/live births, and any physical or behavioral anomalies.

A statistically valid number of animals (males and females) obtained from the F_1 generation and used to produce the F_2 generation should be sacrificed and examined at the appropriate time, with special emphasis on the histopathologic state of the reproductive system. In addition, an adequate number of weanlings of each sex from each dose level, including controls, should also be sacrificed and used for histopathologic analyses.

Data derived from the above should be evaluated for the existence of a relationship between exposure and the incidence and severity of effects on reproduction and behavior, tumors, and mortality. The no-observable-adverse-effect level should also be determined.

Compared to a teratology investigation, which requires about 3 months to complete the exposure and to analyze the results, a three-generation reproduction study represents a considerably longer expenditure of time, approaching 1.5 years between the initiation of the study and the completion of the analysis of the gathered data. However, an elementary profile on teratogenicity and reproductive performance can be obtained within a period of a year or less as a part of a continuing long-term toxicity study by adding the appropriate number of animals at the beginning of the long-term study (2).

Metabolism

Metabolic assessment studies are not directly used for the assignment of risks, tolerances, or action levels, These studies are used to determine parameters, such as absorption, distribution, storage, and excretion, that may affect the performance of materials in biologic systems, thus enabling the researcher to design testing protocols that measure the overall effect of exposure to the material rather than just measuring a part of the biologic response. These protocols can then take into account such factors as tissue concentration, length of time in contact with specific organs or tissues, ease of reactivity, and the production of significant (or nonsignificant) changes

in overall body concentrations that may alter the outcome of a specific testing regimen.

Metabolic study systems are usually centered around the concept that the circulatory system is the major means of transportation of the material, regardless of the route of exposure, Although specific materials may be readily metabolized at or near the site of entry into the body (e.g., the lungs, skin, and intestines), the majority of materials are transported unchanged to the liver and then distributed via the circulatory system to other tissues and organs. At the same time, some of the material may be excreted unchanged in the urine, feces, and air or converted to various metabolizes that are excreted or bound in the tissues. The proportions of the metabolic products that are excreted or bound depends on the chemical nature of the original compound, the dose, route of administration, species, strain, sex, diet, and environmental factors.

The objective of the metabolic study is to mathematically evaluate the rates and relative importance of these processes in limiting the concentration of materials in the tissues of the body. For this purpose, the body is usually visualized as a group of pharmacokinetic compartments, a simplistic view that surprisingly approximates fairly accurately the complex, interdependent processes that actually occur in the body (95). These compartmental models allow the researcher to use measurements of blood concentration as indirect estimates of tissue concentrations and to determine the length of time the material remains unaltered (i. e., the biologic half-life). Rates of absorption from one compartment to another can also be readily measured, thus giving the researcher information on the rate constants of diffusion into tissues and on the rates of distribution from the blood to various tissues. Clearance or elimination rates can also be determined for the excretion of the material from the body via the kidneys, gastrointestinal tract, lungs, saliva, and perspiration.

The rate of metabolism of materials depends on many factors, among the most important of which are the physiochemical characteristics of the molecule itself. Polar compounds are usually excreted very rapidly from biologic systems largely unchanged because of the chemical activity of these compounds, whereas nonpolar compounds such as lipids usually remain in the body longer because they must be metabolized to polar compounds before excretion occurs.

Another important factor in the metabolism of a compound is its structural resemblance to naturally occurring substances in the body. Foreign compounds that closely resemble normal body constituents are frequently metabolized by the same specific enzyme systems that metabolize their normally occurring analogues. Most foreign compounds, however, have no endogenous counterpart and must be metabolized by relatively nonspecific enzyme systems. These nonspecific enzymes catalyze many different types of reactions leading to a diversity of metabolic products. In general, the nonspecific enzyme reactions can be categorized into two types, The first includes the conversion of one functional group into another (oxidation of alcohol to aldehyde), the splitting of neutral compounds to fragments having polar groups (hydrolysis of esters and amides), or the introduction of polar groups into nonpolar compounds (hydroxylation). The second type includes the conjugation of the created polar group with glucuronate, sulfate, glutathione, or methyl groups to form a soluble, excretable product. The product formed in the first reaction may be either more or less toxic than the parent compound or may possess a different type of toxicity. Often, it is this product that actually causes the toxic effects, including cancer, mutations, cellular necrosis, hypersensitivity, fetotoxicity, and blood dyscrasias. A portion of the chemically reactive metabolize formed becomes bound to tissue macromolecules, such as cellular proteins, DNA, RNA, glycogens, or lipids. In this way, it disrupts the normal function of the macromolecule, causing adverse biologic effects. In contrast, the majority of the second type of reaction products are usually either nontoxic or considerably less toxic than the parent compound,

The metabolism of a foreign material is controlled by enzymes, and any factor which affects these enzymes also affects the metabolism of the compound and consequently its toxicity. The metabolism of a compound may be inhibited or stimulated by the presence of competing substrates. Pretreatment with drugs, steroids, food additives, pesticides, polycyclic hydrocarbons, polycyclic amines, and normal constituents of food can result in an increase in the activity of enzymes that metabolize foreign compounds. This increase in enzyme activity differs according to the inducing chemical, but is mediated through an increased rate of synthesis of enzyme protein, a decreased rate of turnover of enzyme, or an activation of enzyme, possibly by changes of structure or conformation. Because of this phenomenon, chronic administration of a foreign material may enhance

the activity of the enzymes that catalyze its metabolism.

Several types of biologic phenomena are readily studied using a metabolic test system. These include activation, antagonism, synergism, and potentiation. Activation has been briefly touched on above with the description of the process of toxification: it involves the rendering of an inactive molecule into an active molecule, usually through the removal or substitution of a neutralizing factor attached to the molecule, but it may also involve the direct addition of a constituent to the molecule. One of the most common examples of activation, the conversion of a precarcinogen to a carcinogen, is the addition of oxygen molecules to polycyclic aromatic hydrocarbons via the microsomal NADPH-dependent cytochrome P-450 mixed function oxidase enzyme system. This conversion to the oxide produces a carcinogenic agent of considerable potency, whereas the parent material does not possess a direct carcinogenie potential per se. The determination of metabolic activation can permit the experienced researcher to predict the kind of adverse effect that is likely to be elicited from the parent molecule, thus enabling the researcher to better design experiments to observe these effects.

From a properly designed and well-carried-out metabolic study, the researcher can gain valuable insight into the potential toxic actions and biologic effects to be expected from a foreign compound. If a compound is not absorbed or if the compound is destroyed or rapidly eliminated from the body there is little likelihood of pronounced toxic effects. If the principle metabolic products are polar, conjugation (and thus elimination) rapidly occurs and again there is little potential for pronounced toxicity. On the other hand, if metabolism produces an activation product, toxicity is enhanced and biologic effects may be pronounced.

APPENDIX C REFERENCES

1. 41 CFR 51206 (1976).
2. Food and Drug Administration Advisory Committee on Protocols for Safety Evaluations: Panel on Reproduction Report on Reproduction Studies in the Safety Evaluation of Food Addit ives and Pest i-icicle Residues (1970). *Tox.Appl.Pharm.* 16:264.
3. *Modern Trends in Toxicology* (1968). E. Boyland and R. Goulding, eds., But terworths, London.
4. Methods *in Toxicology (* 1970). G. E. Pagett, cd., Davis, Philadelphia.
5. Frawlev, J. P. (1955). *Food* Drug *Cosmetic Law J.* 1 o: 703.
6. Smyth, H. F., Jr., and C. P. Carpenter (1944). *J. Ind. Hyg. and Toxicol.* 26:269.
7. Friedenwald, J. S., W. F. Hughes, Jr., and H. Hermann (1944). *Arch. of* Oph *th.* 31:279.
8. Carpenter, C. P., and H. F. Smyth, Jr. (1946). Am. *J. Ophth.* 29: 1363.
9. Draize, J. H., G. Woodard, and H. O. Calvery [1944),]. *Pharm. find Expt. Therapeu.* 82:377.
10. Draize, J. H., and E. A. Kelly (1952). *Proc. of* the *Sci.* Sec. *of Toilet Goods* Assoc. 17:1.
11. Draize, J. H. (1955). *Food* Drug Cosmetics *Law J.* 10:722.
12. Draize, J. H. (1965). *Assoc.* of Food *and* Drug Of *ficials of* the U. S., Topeka, Kans., 46.
13. Lansteiner, K., and J. Jacobs (1935). *J. Exp. Med.* 61: 643.
14. Lansteiner, K., and J. Jacobs (1936). *J. Exp.* Meal, 64:625.
15. Buehler, E. V. (1965). Arch. *Derm.* 91:171.
16. Baer, R. L., S. Rosenthal, and C. J. Sims [1956). *J.* Invest. *Derm.* 27:249.
17. Marzulli, F. N., and H. I. Maibach (1970). *J.* Soc. *Cosmet. Chem.* 21 :695.
18. Maguire, H. C., Jr. (1973). *J.* Soc. *Cosmet. Chem.* 24: 151.
19. Levine, B. B. (1960). *J. Exp. Med.* 112: 1131.
20. Salvin, S. B. (1965). Fed. *Proc.* 24:40.
21. Turk, J. L. (1964). Int. Arch. Allergy 24:191.
22. *Biology of the Immune* Response (1970). P. Abronoff and M. LaVie, eds., McGraw-Hill, N.Y.
23. Textbook of *Immunopathology* (1976). H. G. Muller-Eberhard, cd., Grune and St rat ton, N. Y..
24. *Medical* Physiology (1968). V. B. Mountcastle, cd., Mosby, St. Louis.
25. A Textbook of Histology (1964). W. Bloom and D. W. Fawcett, eds., Saunders, Philadelphia.
26. *Pathology* (1971). W. A. D. Anderson, cd., Mosby, St. Louis.
27. Nelson, A. A. [1955). *Food* Drug Cosmetic *Law J.* 10:732.
28. Comroe, J. H., Jr., R. E. Forster II, A. B. Dubois, W. A. Briscoe, and E. Carleson (1962). *The* Lung: *Clinical* Physiology *and Pulmonary* Func *tion* Tests, Yearbook Medical Publ., Chicago.
29. Armstrong, M. L. (1974). *Electrocardiograms,* Yearbook Medical Publ., Chicago.
30. Current Practice of *Clinical Electroencephalography* (1977). D. W'. Klass and D. D. Daly, eds., Raven.
31. Goodale, R. H. (1965). *Clinical Interpretation of Laboratory Tests,* Davis, Philadelphia.
32. Henry, R. J. (1974). *Clinical Chemistry:* Principles *and* Techniques, Harper Medical.

54-515 0 - 79 - 11

33. Amuino, J.S. (1976). Clinical Chemistry: Principles and Procedures, Little.
34. Xintaras, C., and B. L. Johnson (1976). Essays in Toxicology, W. J. Hayes, cd., 7:155, Academic Press, New York.
35. Ekel, G. J., and W. H. Teichner (1976), A n Analysis and Critique of Behavioral Toxicology in the U. S. S. R., 13 HEW (NIOSH) Publication No. 77-60, Cincinnati.
36. Behavioral Toxicology (1974). C. Xintaras, B. L. Johnson, and I. DeGroot, eds. DHEW (NIOSH) Publication No. **74-126,** Cincinnati.
37. Weiss, B., J. Brozek, H. Hanson, R. C. Leaf, N. K. Mello, and J. M. Spyker (1975). Principles for *Evaluating Chemicals in* the Environment t. 198 National Academy of Sciences, Washington, D.C.
38. McCa arm, J., and B. M. Ames (1976). Ann. N. Y. *Acad. Sci.* 271 :5.
39. Bridges, B. A. (1973). Environ. *Health Perspect* 6:221,
40. Bridges, B.A. (1974). *Mut. Res.* 26:335.
41. Flamm, W. G. (1974). *Mut. Res.* 26:329.
42. Ehrenberg, L. (1973). In: *Chemical* Mu *tagens: Principles and Methods for Their Detection,* vol. 2. A. Hollaender, ed., Plenum Press, New York.
43. Underbrink, A. E., L. A. Schaierer, and A. H. Sparrow (1973), In: Chemical Mutagens: Principles and Methods for Their Detection, vol. 2, A. Hollaender, cd., Plenum Press, New York.
44. Ames, B. N.. F. D. Lee, and W. E. Durston (1973). *Proc. Nat. Acad. Sci.,* U.S.A. 70:782.
45. Ames, B. N., J. McCann, and E. Yamasaki (1975). Mut. *Res.* 31 :347.
46. Laishes, B. A., and H. F. Stich (1973). *Biochem. Biophys. Res. Commun.* 52:827.
47. Brusick, D. J., and V. W. Mayer (1973). *Env. Health Perspect.* 6:83.
48. Ong, T-M, and H. V. Malling (1975). Mut. *Res.* 31: 195.
49. Mohn, G., J. Ellenberger, and D. McGregor (1974). *Mut. Res.* 25: 187.
50. Mortimer, R. K., and T. R. Manney (1971). In *Chemical* Mutagens: Principles *and* Methods *for* Their *Detection,* vol. 1, A. Hollaender, cd., Plenum Press, New York.
51. DeSernes, F. J., and H. V. Malling (1971). In: *Chemical* MU *tagens,* Principles *and Methods for* Their *Detection,* vol. 1, A. Hollaender, cd., Plenum Press, New York.
52. Zimmerman F.K. (1975). Mut *Res.* 3 1-71.
53. Gabridge, M. G., A. DeNunzio, and M. S. Legator (1969). N(ltur[?68:221,
54. Siebert, D. (1973). Mut. Res. *17:307.*
55. *Durston, W. E., and B. N. Ames (1974). Proc. Nat. Acad. Sci.* USA 71 :737.
56. Commoner, B., A. J. Vithayathil, and J. Henry (1974). *Nature* 249:850.
57. Brew en, J. G., P. Nettesheim, and K. P. Jones (1970), *Mut. Res.* 10:645.
58. Legator, M. S., and H. V. hi;} lling [1971). In: *Chemical Mutagens: Principles and Methods* for Their Detection, vol. 1, A. Hollaender, cd., Plenum Press, New York,
59. Slater, E. E., M. D. Anderson, and H. S. Rosenkranz (1971). *Cancer Res.* 31:970.
60. Chu, E. H. Y. (1971). In: *Chemical* Mutagens: Principles *and* Methods *for* Their Detection, vol. 1, A. Hollaender, ed., Plenum Press, New York.
61. Kao, F-T, and T. T. Puck (1974). Methods *Cell Biol* 8:23.
62. Cline, D., and J-A F. S. Spector (1975). Mut. *Res.* 31: 17.
63. Regan, J. D., **and** R. B. Setlow (1973). In: Chemical *Mutagens:* Principles and Methods *for* Their Detection, vol. 2, A. Hollaender, cd., Plenum Press, New York.
64. Sega, G, A., J. G. Owens, and R. B. Cumming (1976). Mut. *Res.* 36: 193.
65. Taylor, J. H. (1958). Genetics 43:5 15.
66. Allen, J. W., and S. A. Latt (1976). Nature 260:449.
67. Abrahamson, S., and E. B. Lewis (1971). In: Chemical Mutagens: Principles *and* Methods *for* Their Detection. vol. 1, A. Hollaender, ed., Plenum Press, New York.
68. Vogel, E., and B. Leigh (1975), Mut. *Res.* 29:383.
69. Casida, J. E. (1969). In: *Microsomes and* Drug *Oxidation,* J. R. Fouts and G. J. Mannering, eds., Academic Press, New York.
70. Smith, R. H., and R. C. von Borstel (1973). In: *Chemical Mutagens:* Principles *and* Methods for *Their* Detection. vol. 2, A. Hollaender, ed., Plenum Press, New York.
71. Schmid, W. (1975). *Mut. Res.* 31 :9.
72. Leonard, A. (1973). In: Chemical *Mutagens:* Principles *and Methods for* Their Deter; *tion,* vol. 2, A. Hollaender, cd., Plenum Press, New York.
73. Leonard, A. (1975). *Mut. Res.* 31 :291.
74. Russell, 1,. B. (1976). [n: *Chemical Mutagens:* Principles *and Methods for Their Detection,* vol. 4, A. Hollaender, cd., Plenum Press, New York.
75. Bateman, A. J., and S. S. Epstein (1971). In: *Chemical Mutagens:* Principles *and Methods* for Their Detection, vol. 1, A. Hollaender, cd., Plenum Press, New York.
76. Sontag, J. M., M. P. Page, and U. Saffiotti (1976). In: *Guidelines* for *Carcinogen Bioassay in Small Rodents.* U.S. Department of Health, Education and Welfare, Public Health Service, National Institutes of Health. National Technical In formation Service PB-264 061, Springfield, Va.
77. Saffiotti, U., and H. Autrup (1978). In: *In* Vitro *Carcinogenesis: Guide to* the *Literature, Advances and Laboratory* Procedures, U.S. Department of Heal t h, Education and Welfare, Public Health Service, National Institutes of Health. National Technical Information Service PB-281 737, Springfield, Va.
78. Carcinogenicity (1973). In: The Testing *of Chemicals for Carcinogenicity,* Mu *tagenicity, and Teratogenicity,* Ministry of Health and Welfare, Canada.

79. 43 CFR 37336 (1978).
80. Purchase, I. F. H., E. Longstaff, J. Ashby, J. A. Styles, D. Anderson, P. A. Lefevre, and F, R. Westwood (1978). *Br. J. Cancer* (in press).
81. St vles, J. A. (1977), *Br. J. Cancer* 36:558.
82. Williams, I). J., and B. R. Rabin (1971). *Nature* 232: 102.
83. Bock; F. H., and R. Mund (1958), *Cancer Research* 18:887.
84. Iversen, O. H., and A. Evensen (1962). In: *Experimen tal* Skin *Ca* rein *ogenesis* in *Mice.* Norwegian Universit y Press.
85. Longstaff, E., and F. R. Westwood (1978). Unpublished.
86. DiPaola, J. A., R. L. Nelson, P. J. Donovan, and C. H. Evans (1973). *Archs. Pathol.* 95:380.
87. Epstein, J., R. W. Rosenthal, and R. J. Ess (1955). *Anal. Chem.* 27: 1435.
88. Szent-Gyorg yi, A., 1. Isenberg, and S. L. Baird (1960). *Proc. Nat. Acad. Sci. U.S.A.* 16: 1444.
89. Szent-Gyorgyi, A., and J. McLaughlin (1!361). *Proc. Nat. Acad. Sci.* 11. S.A. 47: 1397.
90. McCann, J., and B. N . Ames (1976). *Proc. Nat. Acad. Sci. U.S.A.* 73:950.
91. Heddle, J. A., and W . R. Bruce (1977). In: *Origins of Human Cancer. Book* C. *Human* Risk Assessment. H. H. Hiatt, J. D. Watson, and J. A. Winsten, eds. Cold Spring Harbor Labor a torv.
92. Teratogenicity (1973). In: The *Testing* of *Chemicals for Carcinogenicity*, M u *tagenici* ty, *and* Te I*togenicity*. Ministry of Health and Welfare, Canada.
93. Courtney, K. D., and M. Chernoff (1974). In: *Training Manual for Tera tology*. Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, N.C.
94. Green, E. L. (1962). *Genetics.* N.Y. 47: 1085.
95. Compartment ts, *Pools and Spaces* i n *Medical Physiology* (1967). P.-Ii. E. Bergner and C. C. Lushbaugh, eds. U.S. Atomic Energy Commission, Division of Technical In forma t ion, Oa k Ridge, Tenn.
96. Purchase, I. (1976). *Nat ure* 264:624.