SAFETY EVALUATION AND RISK ASSESSMENT USING IMMUNOTOXICOLOGY METHODS

Jack H. Dean and Jeff Hincks
Sanofi Research, Sanofi Pharmaceuticals, Inc., Malvern, Pennsylvania, USA

Michael I. Luster
National Institute of Occupational Safety and Health, Morgantown, West Virginia, USA

G. Frank Gerberick
Procter and Gamble Co., Cincinnati, Ohio, USA

David A. Neumann
ILSI Risk Science Institute, Washington, DC, USA

Kenneth L. Hastings
Center for Drug Evaluation and Research, Division of Antiviral Drug Products, U.S. Food and Drug Administration, Rockville, Maryland, USA

BACKGROUND (J. DEAN)*

This article represents a composite of ideas expressed by the authors in a symposium, Safety Evaluation and Risk Assessment Using Immunotoxicology Methods, presented at the 1996 Annual Meeting of the American College of Toxicology. The purpose of this symposium was to examine the current status of immunotoxicity assessment methodology and its application in the safety evaluation and risk assessment of drugs and chemicals.

Chemical- and drug-induced effects on the immune system are thought to be manifested by either suppression or induction of aberrant immune responses, such as hypersensitivity or allergy (National Research Council, 1992). Evidence of the immunosuppressive effects of chemicals in rodents is well established (Urso and Gengozian, 1980; Holsapple et al., 1983; Pestka et al., 1987; see review by Dean and Murray, 1991) and in humans (Bekesi et al., 1978; Lew et al., 1986; Boyer, 1989) is well documented. Experimental procedures for the assessment

*Name in parentheses indicates section was contributed by and reflects the opinion of this author.

Received 4 June 1997; accepted 22 August 1997.

This article was co-authored by Dr. Hastings in his private capacity. No official support or endorsement by the Food and Drug Administration is intended or should be inferred.

Address correspondence to Jack H. Dean, PhD, President and Scientific Director, Sanofi Research, 9 Great Valley Parkway, P.O. Box 3026, Malvern, PA 19355, USA.

1091-5818/98 $12.00 + .00

of the potential of chemicals to induce immunosuppressive (including myelosuppressive) effects in rodents are well established and validated (Luster et al., 1992, 1993). Compound development and marketing decisions, when warranted, are often based on preclinical evidence that a compound has immunosuppressive effects.

Procedures for assessing hypersensitivity in chemical exposed rodents are routinely used in safety evaluation studies (Karol et al., 1985; Buehler, 1985), and various compounds have been regulated on the basis of their associated hypersensitivity reactions (Office of Technology Assessment, 1991).

This article focuses on the current status of immunotoxicity testing methods and newer approaches for assessing immunosuppression, myelosuppression, hypersensitivity, and autoimmunity. It examines the practical application of these methods in safety evaluation and risk assessment relative to the extent of their characterization and predictive value. Newer methodology is described that holds promise for the future and may further enhance risk assessment. Finally, we explore strategies in immunotoxicity assessment that might aid in the development of a new drug candidate.

The term "risk assessment" is widely used in this article and deserves some definition. Risk assessment was best defined by an often cited National Academy of Science report (1983), and is said to be comprised of four steps: hazard identification, dose-response assessment, exposure assessment, and risk characterization. While exceptions exist, the majority of data from immunotoxicity studies have not been widely or routinely used in this process. Although direct comparisons of human clinical and animal study data are limited, other factors may also be responsible for the minimal use of these data in risk assessment. This may include a concern that immunotoxicity testing has often been conducted without full knowledge of its predictive value in humans or its quantitative relationship to immune-mediated diseases. Increasing numbers of immunotoxicology studies in human populations have demonstrated subtle immune test changes including altered lymphocyte phenotype distribution. While these immune changes represent dose-related, statistically significant effects, the data are usually within the normal ranges and are not associated readily with detectable disease. Thus, establishing the quantitative relationship between immune parameter changes and disease progression is paramount for accurate risk assessment.

IMMUNOSUPPRESSION (M. LUSTER)

Luster et al. (1988) have previously reported on the design and content of a screening test battery involving a "tiered" approach (Table 1) for

Table 1. Panel recommended for detecting immunomodulation following chemical or drug exposure in rodents from the NIEHS/NTP (Luster et al., 1988)

Parameter	Procedures
Screen (Tier I)[a]	
Immunopathology	Hematology: complete blood count and differential Weights: body, spleen, thymus, lymph node, kidney, liver Cellularity: spleen, bone marrow Histology: spleen, thymus, lymphoid node Quantitation of splenic B, T, and lymphocyte subpopulations
Humoral immunity	Enumerate immunoglobulin (Ig) M antibody to T-dependent antigen (sRBC)
Nonspecific immunity	Natural killer (NK) cell activity
Comprehensive (Tier II)	
Nonspecific immunity	Macrophage function—quantitation of resident peritoneal cells and phagocytic ability (basal and activated by MAF) RES clearance
Humoral immunity	Secondary (IgG) antibody response
Host resistance challenge models[b]	Syngeneic tumor cells PYB6 sarcoma (tumor incidence) B16F10 melanoma (Lung burden) Bacterial models: *Listeria monocytogenes*, Streptococcus sp. Viral models: influenza Parasite models: *Plasmodium yoelii*

[a] The testing panel was developed using B6C3F1 female mice, but have been successfully adapted for use in rats.

[b] For any particular chemical tested, only two or three host resistance models are selected for examination.

detecting potential immunosuppressive compounds in mice as part of a National Toxicology Program development and validation effort. This test battery has been used to examine a variety of compounds, and the databases generated from these studies include an expanded dose study with the classical immunosuppressant cyclophosphamide. The National Toxicology Program (NTP) database, consisting of over 50 compounds, has been analyzed in an attempt to improve the accuracy and efficiency

of screening chemicals for immunosuppression and to identify tests that better predict experimental-induced, immune-mediated diseases (Luster et al., 1992, 1993). While a number of limitations existed in the analyses, several conclusions were obtained from the results:

1. Normally, examination of only two or three immune parameters is needed to successfully identify an immunotoxicant. In particular, lymphocyte subpopulation enumeration and quantification of the antibody response to T-cell dependent antigens appeared particularly beneficial. Furthermore, some commonly measured parameters (e.g., leukocyte counts, lymphoid organ weights) were found to be fairly insensitive.
2. A good correlation was found between changes in the immune tests and altered host resistance, in that there were no instances of altered host resistance without significant changes in the immune test parameter(s). However, in many instances, immune changes were observed in the absence of detectable changes in host resistance. This was interpreted as a reflection that immune tests are generally more sensitive than the host resistance assays.
3. Several assays were relatively good indicators, while several, such as leukocyte counts and proliferative response to lipopolysaccharide (LPS), were poor indicators for host susceptibility changes. Some of the tests that gave the highest association with host challenge models (tumor cell or infectious agent challenge) were those that we described previously as being the best indicators for immunotoxicity, such as the antibody plaque assay and surface marker analysis. The plaque forming cell (PFC) and surface marker tests have individual predictive values for altered host resistance of 78% and 83%, respectively, and could reasonably be described as highly predictive stand-alone tests. The only paired tests that exceed 83% joint (pairwise) predictive value were pairings of these assays (PFC or surface antigens) with other tests, and for many such pairings the predictive values increased only minimally. Thus, the PFC and, especially, surface marker tests have high individual predictive values for immunosuppression.
4. Considering that a "background" level of infectious diseases exists in a population, it is possible that moderate changes in immune function would translate to a clinically significant change in host resistance outcomes, given that the population exposed is sufficiently large. While this would be difficult to demonstrate in the human population, it can be demonstrated experimentally in animals using "infectivity models" whereby the virulence of the agent or challenge dose of the infectious agent is controlled. In using one of these challenge models (see Figure 1), it was observed that tumors will develop even in control mice provided sufficient numbers of tumor cells to overcome natural resistance are administered.

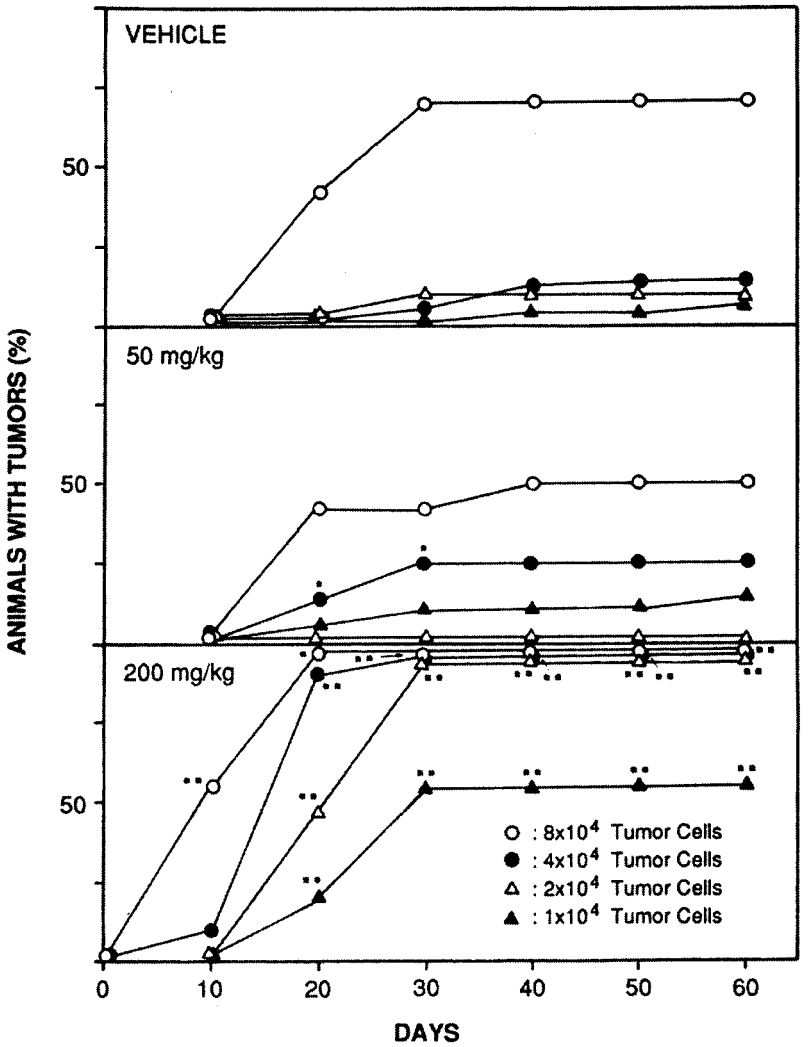

Figure 1. Effects of cyclophosphamide on PYB6 tumor formation as a function of tumor cell challenge. Groups of mice were administered a single ip injection of saline or 50 or 200 mg/kg cyclophosphamide, followed 48 h later by subcutaneous injection of either 1×10^4, 2×10^4, 4×10^4, or 8×10^4 PYB6 tumor cells. The data are presented as the percent of animals that develop palpable tumors within 60 days. Each point represents 20 mice. Asterisk indicates significant, $p < .05$, and double asterisk, $p < .01$, versus vehicle-treated group by Fisher's exact test. Reprinted from Luster et al. 1993. As shown in the top panel, groups of animals with a "normal" immune system can develop increased frequency of tumors, given that the number of tumor cells administered is high. In contrast, animals receiving the 200-mg/kg dose of cyclophosphamide (bottom panel) all develop tumors, albeit at different frequencies. In animals mildly immunosuppressed (50-mg/kg dose level, center panel), small differences in tumor frequency can be detected depending on the number of tumor cells administered.

5. Regression modeling, using the cyclophosphamide data, indicated that most, but not all. of the immune function–host resistance relationships follow a linear model. When the data from all 50 chemicals were combined, this trend was evident; however, it was not possible to statistically confirm linear or threshold relationships because of the limited number of doses. Thus, a more mechanistically based mathematical model will have to be developed.
6. Finally, using one data set, methods were developed for modeling quantitative relationships between changes in selected immune assays and host resistance tests. It is impossible, at present, to determine how applicable these analyses will be for immunotoxic compounds with different immune profiles. However, as more analyses become available, our ability to accurately estimate potential clinical effects from immunological tests in animals should improve.

MYELOSUPPRESSION (J. HINCKS)

Myelosuppression is often one of the dose-limiting toxicities associated with anticancer and anti-AIDS (AIDS-acquired immune deficiency syndrome) agents. Thus, the assessment of human myelotoxic potential of new chemical entities, prior to clinical trials, is of critical importance in the process of drug development. This type of assessment during drug discovery can assist in the selection of candidate compounds for further development as well as assist clinicians in the selection of the maximum tolerated dose in humans. Classical preclinical toxicology studies in mice, rats, and dogs provide an indication of the degree of drug-induced bone-marrow suppression and recovery. However, it is difficult to relate myelosuppression detected in animal studies to a possible human response without some human data. In vitro bone-marrow stem-cell assays, together with in vivo preclinical toxicology studies, can be used to predict the human response prior to clinical trials (see Figure 2). This prediction paradigm is based on determining species differences between mouse or dog and human using in vitro myelotoxicity assays, coupled with careful evaluation of the in vivo preclinical toxicology results. Different types of in vitro bone-marrow stem-cell assays can be conducted to provide myelotoxicity information on different hematopoietic lineages of the bone marrow. For example, methodologies for in vitro bone-marrow stem-cell assays are available to detect colony-forming units (CFUs) in the erythroid, granulocyte and macrophage, and megakaryocytic series (CFU-E, CFU-GM, and CFU-Meg, respectively).

For a general review of the culture conditions for in vitro bone-marrow stem-cell assays see Harvey et al. (1984). These types of assays have been used in the past (Du et al., 1990) to assess the in vitro myelotoxicity of anti-HIV agents such as 3'-azido-3'-deoxythumidine (AZT), dideoxyinosine (ddI), and dideoxycytosine (ddC). In addition, in vitro

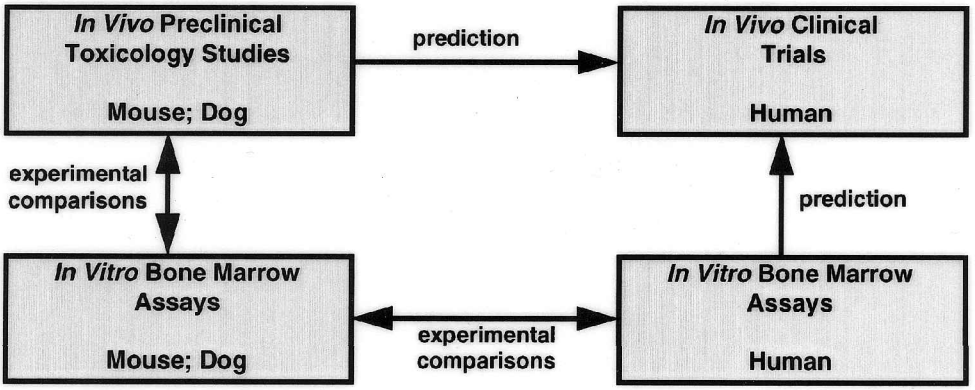

Figure 2. Experimental design using in vitro myelotoxicity studies and in vivo preclinical toxicology study results to predict human myelosuppression of new drugs prior to Phase I clinical trials. From J. Hincks.

bone-marrow stem-cell assays have been used to assess the relative efficacy and toxicity of two antileukemia drugs, busulfan and hepsulfam (Hincks et al., 1990). Methods have been developed to support the growth of CFU-E, BFU-E (blast-forming units, erythroid), and CFU-GM from mouse, dog, and human bone marrow in a methylcellulose support matrix. The mouse and dog assays were selected by Hincks et al. (1990) based on the fact that these are the most common species used in preclinical toxicology studies for anticancer drug evaluation. The human assay was developed in one of our laboratories (J. Hincks) in order to correlate preclinical and clinical responses (see Figure 2), and then was applied to assess myelotoxicity in human bone-marrow stem cells of new anticancer drugs prior to clinical trials. Two new chemical entities, SR259075 and SR333377, were tested in this system (Fischer and Hincks, 1992; Fischer et al., 1995). Results of these studies indicate the following utility of in vitro bone marrow assays:

1. Can provide human myelotoxicity data prior to Phase I clinical trials.
2. Can assist clinicians in defining a safe starting dose for Phase I clinical trials.
3. Can assist clinicians in designing safe and rapid dose escalation strategy.
4. Can reduce new drug development time and resources.

CONTACT HYPERSENSITIVITY METHODS (G. F. GERBERICK)

Various classical guinea pig tests have been developed for evaluation of skin-sensitizing potential of chemicals or drugs. Among those most rigorously validated and widely applied are the guinea pig maximization test (GPMT) (Magnusson and Kligman, 1969) and the Buehler occluded

patch test (Buehler, 1965, 1985). Thus, the current accepted animal testing methods for predictive assessment of contact sensitization are limited to various guinea pig skin sensitization protocols that measure cutaneous erythema and/or edema in induced, sensitized, and challenged animals (Botham et al., 1991). A newer approach is represented by the murine local lymph node assay (LLNA), which represents a predictive test that utilizes in vivo cell proliferation in the draining lymph nodes for assessment of the contact sensitization potential or materials (Kimber et al., 1986, 1989). This assay measures the proliferative response in the draining auricular lymph nodes during the induction phase of a contact sensitization response after topical exposure to materials. It is based on the observation that the induction phase of a contact sensitization response to allergen is characterized by lymphocyte proliferation and hyperplasia in the lymph nodes draining the site of topical exposure. The LLNA has been proposed for use in screening materials for their contact sensitization potential.

Validation Status of Murine LLNA

To define the role of the LLNA in predictive testing, results from the assay have been compared with results from guinea pig and human tests. Specifically, the LLNA assay has been evaluated extensively for its reliability and sensitivity in predicting the contact sensitization potential of materials through both internal and interlaboratory validation studies (Loveless et al., 1996; Kimber et al., 1995; Kimber and Basketter, 1992; Basketter et al., 1991) and has been found to correlate well with guinea pig studies and analyses in humans (Basketter et al., 1993, 1994). For example, Kimber et al. (1990) reported a comparative analysis in which 24 chemicals, of previously unknown contact sensitizing potential, were evaluated in both the local lymph node assay and the occluded patch test of Buehler. The data reported demonstrate that the local lymph node assay successfully identified those chemicals that were classified as moderate or strong skin sensitizers in the Buehler test. Basketter et al. (1991) evaluated the performance of the LLNA with 25 chemicals on which guinea pig maximization test or Buehler occluded patch test data were available. The 25 chemicals included preservatives, perfume ingredients, surfactants, plastics/resin chemicals, and oil additives. A high level of agreement between the results of local lymph node assays and guinea pig test data was found. Recently, the LLNA was evaluated in two international studies by five independent laboratories. In the first international study, two sensitizers and one nonsensitizer were evaluated (Kimber et al., 1995). All five laboratories identified the two known sensitizers as being positive in the LLNA. In the second international study, the sensitivity and selectivity

of the assay were examined further by analysis of six additional chemicals. All five laboratories identified as positive the five moderate to strong sensitizers and the one nonsensitizer as negative. However, sodium lauryl sulfate, considered to be a nonsensitizing skin irritant, was identified as positive in the assay. The overall conclusion from these international studies is that five laboratories successfully and consistently employed the LLNA to reach identical conclusions on the sensitizing potential of nine chemicals. Finally, in a comprehensive review of published and unpublished LLNA data, a high level of concordance between LLNA and guinea pig data was reported for a wide range of chemicals (Basketter et al., 1996).

As stated, an essential comparison for the LLNA is with human data. Basketter et al. (1994) compared human maximization test results with those obtained with the LLNA. The human maximization test method is a rigorous assessment of the sensitization potential of chemicals in humans. The authors reported that the LLNA identifies those chemicals that are significant human contact allergens and that the specificity of the assay is good.

Advantages and Limitations of the Murine LLNA

As the skin sensitization potential of chemicals has been determined by using guinea pig tests, there is a strong temptation to judge the merits of a new assay validated against the "gold standard" of a guinea pig assay such as the maximization test or Buehler test. As discussed earlier, the LLNA shows good concordance with guinea pig test results. However, technical aspects of test conduct and species differences between guinea pigs and humans inevitably mean that no such gold standard exists. Seventy-seven chemicals were tested with results indicating 89.6% concordance between studies. Guinea pig or mouse data may not always mirror precisely and quantitatively the extent of the hazard to humans. In general, it is difficult to compare predictive tests with human data, since the former relate to hazard identification whereas the latter represent problems arising from the expression of that hazard: risk. Fortunately, the LLNA assay is capable of identifying chemicals that have the potential to cause significant allergic contact dermatitis in humans.

Compared with classical guinea pig methods, the LLNA offers a number of important advantages, not the least of which are that:

1. The endpoint is both quantitative and objective and not subject to interpretative difficulties associated with the visual assessment of challenge-induced skin reactions.

2. It offers the advantages of speed and comparatively low cost (and thus the opportunity for high throughput).
3. It uses about half the number of animals required for organization for Economic Cooperation and Development (OECD) protocol GPMT and Buehler tests.
4. It also offers a clear advantage for evaluation of colored compounds and dye chemicals which may obscure cutaneous erythema (Kimber et al., 1994).
5. There is a further animal welfare advantage in that the LLNA does not require the use of Freund's compete adjuvant, intradermal injections of test substance, fur removal, occlusive dressings, or the use of restraints, which are all features of the GPMT and/or the Buehler test. Thus, there is an important opportunity for both substantial reduction and refinement of animal usage in a manner that will not compromise the existing standard of classification and labeling of skin sensitization in the European Union.

Chemicals that are classified as having moderate or greater sensitizing potential on the basis of guinea pig tests and many that are considered weak or mild sensitizers prove positive in the LLNA. Unfortunately, the local lymph node assay fails to identify as positive some chemicals that provoke comparatively weak responses in the more sensitive guinea pig tests. However, as defined in the classification thresholds set by the European Economic Community, these chemicals need not be positive in a suitable alternative assay.

The LLNA may also be of value for the assessment of irritant chemicals. It is a requirement of guinea pig tests, which rely upon measurement of induced erythema in previously sensitized animals, that challenge is performed with a concentration of the test material that is unable to provoke skin irritation or inflammation in nonsensitized controls. It is possible with highly irritating materials that the concentrations selected for challenge are below those necessary to elicit a contact hypersensitivity reaction. In contrast, the use of high concentrations of irritants in the LLNA can be problematic in some instances. For example, sodium lauryl sulfate (SLS), benzoylkonium chloride, and salicylic acid are a few irritants that can give positive responses in the LLNA (Montelius et al., 1994; Gerberick et al., 1992). Thus, the question is, why do irritants cause positive responses in LLNA tests? One explanation could be that irritants cause lymph-node proliferation by a nonimmune mechanism involving the stimulation of epidermal cytokines. Clarification using an alternative procedure, or by more detailed analysis of the lymph-node response, may be indicated. Potential endpoints for use in differentiating irritant and allergic responses include quantification of B and T cells as well as analysis of T-cell activation markers on

CD4 and CD8 cells (Gerberick et al., 1996; Sikorski et al., 1996). It is very important to point out, however, that not all irritants are positive in the LLNA (Kimber et al., 1994; Gerberick et al., 1992).

Use of Murine LLNA in Safety Evaluation and Risk Assessment

The LLNA is already named in the OECD *Guidelines for the Testing of Chemicals*, Volume 1, *Skin Sensitization*, where it is presented as a screening method. The guidelines state that when the LLNA result is negative it is necessary to conduct a confirmatory guinea pig test. However, the data available on the LLNA (Basketter et al., 1996) strongly support the view that the assay detects "most important contact allergens." Moreover, it may be concluded from the data that a substance negative in the LLNA may not be considered an important contact allergen for humans and that no further animal testing may be necessary.

In conclusion, it can be stated that the LLNA offers a number of advantages over available guinea pig methods. The assay is rapid, relatively cost-effective, and requires small amounts of test material. Exposure is via the relevant route and there is no requirement for adjuvant. The endpoint is objective and quantitative and appears to be largely uninfluenced by the color of the test chemical. The LLNA has been evaluated as a methodology for the identification of significant skin sensitizers and is accepted by some regulatory agencies as a method suitable for the classification of the skin-sensitizing potential of chemicals.

SYSTEMIC HYPERSENSITIVITY AND AUTOIMMUNITY METHODS (D. NEUMANN)

Although the notion that chemical exposure might result in the induction or modulation of systemic or general hypersensitivity or autoimmunity is generally accepted, evidence for such effects is limited in animals (Bigazzi, 1988; National Research Council, 1992). In part, this reflects uncertainty about how to distinguish between normal self-reactivity and autoimmune disease and how to prospectively evaluate chemicals for their potential to elicit or modulate autoimmune responses. Consequently, there is little evidence that procedures for assessing the hypersensitive or autoimmune status of rodents in preclinical studies are sufficiently robust to yield data on which to base development, regulatory, or marketing decisions.

Rodents as well as humans constitutively express self-reactive effectors, that is, autoantibodies and self-reactive T cells (e.g., see Shoenfeld and Isenberg, 1989). Background concentrations of such effectors are

not well established, and criteria for distinguishing between background levels and those associated with pathogenesis or other adverse health effects are elusive, except in specific clinical settings (Condemi, 1992) or in association with certain animal models of autoimmunity (e.g., see Cohen and Miller, 1994). From both theoretical and pragmatic perspectives, the distinction between background autoimmunity and disease-associated autoimmune status often becomes mired in uncertainty about clonal anergy of B and T lymphocytes and the induction and loss of self-tolerance (Abbas et al., 1991). Although these processes are reasonably well understood, it is unclear if such processes can be well addressed in immunotoxicity testing.

Another confounding issue, when considering design of studies to determine whether exposure to a chemical elicits an autoimmune response, is uncertainty about the relevant autoantigen. Virtually any self molecule in any organ or tissue may be the target of immune self-reactivity. In addition, the induction of self-reactivity may be associated with the release of sequestered self-antigens, expression of altered self-antigens, self-antigens bearing foreign antigen determinants, idiotypic antigens, aberrant expression of self-antigens, polyclonal activation of self-reactive cells, regulatory dysfunction, or antigenic mimicry (Shoenfeld and Isenberg, 1989; Rose and Mackay, 1992). The inability to predict the relevant autoantigens on which to base an autoimmune toxicologic assessment is an impediment to attempting such studies.

A potentially useful concept for approaching autoimmunotoxicity in the context of preclinical studies is that of examining for biomarkers of effect (National Research Council, 1992). This concept suggests that exposure to a chemical that can elicit or modulate autoimmunity would result in a demonstrable effect on autoimmune effectors. Detection of such effects does not require a prior knowledge of the specificity of the autoimmune effectors, the specific autoantigenic target, and mechanisms of induction and loss of self-tolerance or the process(es) leading to the observed effect. Rather, by using tissue-based autoimmune effector detection methods, investigators can screen for autoimmunotoxic effects.

Autoimmune-mediated pathologies often are associated with target tissue inflammation, suggesting that a useful starting point would be a thorough examination of animals on test for gross abnormalities with possible inflammatory involvement. Key signs might include fibrotic changes, hemorrhage, or anatomic changes in tissue/organ appearance, such as enlargement. Routine hematoxylin and eosin staining of suspect abnormalities can confirm the presence of inflammatory infiltrates and reveal information about the phenotypes of infiltrating cells. Direct immunohistochemical staining of inflamed tissue can establish the phenotypes and activation status of the inflammatory cells, determine whether immunoglobulin is deposited in the lesion, establish whether

complement plays a role in the observed pathology, and assess whether major histocompatibility complex (MHC) molecule expression has been upregulated in such a manner as to suggest autoimmune activation. Such procedures are typically most effective with sections of frozen tissue (which preserves antigen expression, although paraffin-embedded tissues may be used with some staining reagents) and require careful attention to the use of control tissues and reagents (Tubbs et al., 1986). Although immunohistochemical analysis is often qualitative, both semiquantitative scoring and quantitative procedures, such as quantitative microscopy, image analysis, and radiolabeled secondary antibodies, have been described (Kuper et al., 1995).

Indirect immunohistochemical procedures can be used to assess for autoantibodies or systemic hypersensitivity by incubating serum from treated animals on frozen sections of perfused tissues from normal or control animals (Wolfgram et al., 1985; Rosenberg et al., 1987). Cell binding assays in which splenic or peripheral blood lymphocytes (or preferably T cells) are incubated on similar frozen tissue sections (Jalkanen and Butcher, 1985; Chin et al., 1990) may yield evidence for T-cell mediated autoimmunity. Because exposure to a test material may result in altered self-antigen expression or self-protein/adduct formation creating novel self antigens (Rose and Mackay, 1992), indirect immunohistochemistry and cell binding assays should also be performed on tissues from treated animals to facilitate detection of self-reactivity to these types of autoantigens.

Newer methods have been proposed that might be useful in some situations (Descotes, 1992) for detecting the potential for systemic hypersensitivity or autoimmunity. One of these, the popliteal lymph node assay (PLNA) is a relatively simple test in which the test article is injected into the foot pad of a mouse and the draining lymph node is then obtained and weighed (Kammuller et al., 1989). Increased relative lymph node weight is an indication of immune stimulation. Taken in the context of known autoimmune effects associated with chemical classes, the PLNA could be valuable. However it is uncertain, at this time, that the PLNA assay would be valuable as a stand-alone assay for assessing chemicals, although validation studies are in progress that should answer this question.

APPROACHES FOR INVESTIGATIONAL NEW DRUGS (K. HASTINGS)

With the exception of skin sensitization studies, immunotoxicology studies per se are rarely conducted by the sponsor on an investigational new drug. The more common practice is for sponsors to determine the possibility that

adverse effects observed in standard repeat-dose toxicology studies could indicate potential immune impairment (Bloom et al., 1987). Inherent in this approach is the notion that the ability to detect potential immunotoxic effects depends on the conduct of standard preclinical toxicology studies. The probability is better that potential immunotoxic effects will be discovered if particular attention is given to directly assessing the immune system. The term "enhanced histopathology" has been coined in Europe to denote a specific approach in which lymphoid tissues (spleen, thymus, lymph nodes) are examined in detail using specific antisera to identify T and B cells to discern changes in resident cell types (Schuurman et al., 1994). It has been proposed that this approach results in an increased potential to discover immunosuppressive effects (Kuper et al., 1995; van Loveren et al., 1996). If evidence of immunosuppressive potential is observed in standard nonclinical toxicology studies, the sponsor could conduct specific follow-up immunotoxicity studies to determine the potential for the investigational drug to impair immune function parameters.

Preclinical studies designed to determine the sensitizing potential of investigational drugs are commonly conducted, especially for those intended to be used topically or via inhalation. The most commonly used guinea pig models are intended to mimic potential clinical effects, such as cutaneous or pulmonary hypersensitivity. However, it should be noted that guinea pig sensitization studies have proven useful primarily in detecting contact sensitizers (Botham et al., 1991). Adaptations designed to detect pulmonary sensitization have not proven particularly useful in detecting relatively weak sensitizers that might produce clinically important adverse effects (Blaikie et al., 1995).

Some commonly used tests for systemic sensitization as the passive cutaneous assay (PCA) and the active systemic assay (ASA) rarely provide clinically relevant results (van der Laan and Dean, 1996). The PCA does detect antigenicity, but, almost always, the compounds giving positive results might also be detected in standard skin sensitization assays (because of inherent reactivity with proteins) or are proteins or other macromolecules that are predictably antigenic. The ASA might be useful in some very unusual situations (especially if metabolism of the drug results in antigenic products) but rarely provides clinically useful information (Verdier et al., 1994).

Considerable attention has been given in recent years to the use of mice in detecting chemical sensitizers. Of particular importance is the mouse local lymph node assay (LLNA) discussed previously. Based on currently available information, the LLNA may be considered satisfactory for preclinical safety evaluation of topical drug products.

Drug-associated autoimmunity is a phenomenon that is rarely if ever, predicted in preclinical toxicology studies. No widely evaluated methods are available to predict the potential of a drug to produce autoimmune

reactions. Adverse reactions in the clinic that mimic hypersensitivity reactions, but do not involve specific immune responses to the drug, are often referred to as pseudo-allergy (Descotes, 1986) and are also not predicted by preclinical studies. Taken in the context of known autoimmune effects associated with drug classes, the popliteal lymph node assay (PLNA) may be valuable once evaluation is completed.

Immunostimulation by a drug candidate is rarely encountered and can be associated with several findings in nonclinical toxicology studies, such as histologic demonstration of lymphocytic infiltrates and nonspecific increases in serum immunoglobulins. For some types of compounds, such as adjuvants, immunostimulation may represent the desired pharmacodynamic effect. The PLNA has been proposed as a general assay for detecting immunostimulants (van der Laan et al., 1997). However, as with autoimmunity, the results of the assay should be evaluated in the context of drug class.

In some situations it may be important to determine if a drug has antigenic potential unrelated to sensitization. This is especially the case in repeat-dose toxicology studies with protein or peptide drugs (Dean et al., 1992). Although demonstrating that a drug is antigenic in a given animal model should not always be taken as a safety concern, this could be important in fully evaluating the results of nonclinical toxicology studies. Antidrug immune responses can alter toxicologic, pharmacodynamic, and especially, pharmacokinetic properties of the drug in animal models. In addition, antidrug immune responses can alter clinical effects or produce adverse effects.

Pseudo-allergic reactions can be modeled in nonclinical toxicology studies and are often predictable based on preclinical findings (Descotes, 1986). Phenomena evaluated include potential for drug-induced histamine release and complement activation. It is especially useful that sponsors determine drug exposure levels associated with pseudo-allergic reactions in order to set safe dose levels in clinical trials. In addition, these findings may contribute to drug labeling.

SUMMARY (J. DEAN)

The field of immunotoxicity has advanced considerably in the past few years with the systematic evaluation of methods to predict immunosuppression, myelosuppression, and certain types of hypersensitivity. International interlaboratory comparisons have been conducted with methods for detecting immunosuppression (White et al., 1994) and contact hypersensitivity with good agreement (Kimber et al., 1995). The standard tier for immunotoxicity assessment, as described by the National Toxicology Program (Luster et al., 1988), and methods suggested by the OECD 407 Guideline (van Loveren et al., 1996) are con-

sidered useful in both the mouse and rat. Likewise, these new methods have been introduced and evaluated in the safety assessment of new pharmaceuticals and chemicals by some companies. Most of the methods developed have been drawn from our experience in clinical immunology and an improved understanding of the immune system through advances in molecular biology and basic immunology. There is still work to be done to better document the correlation between animal immunotoxicity results and adverse effects in humans following exposure to chemicals in our environment, and to define predictive methods for evaluating the potential of a drug to induce allergy and autoimmunity, still a significant unpredicted clinical finding.

REFERENCES

Abbas, A. K., Lichtman, A. H., Pober, J. S. 1991. *Cellular and molecular immunology*, pp. 205–211. Philadelphia: W. B. Saunders.

Basketter, D. A., Scholes, E. W., Kimber, I., Botham, P. A., Hilton, J., Miller, K., Robbins, M. C., Harrison, P. T. C., and Waite, S. J. 1991. Inter-laboratory evaluation of the local lymph node assay with 25 chemicals and comparison with guinea pig test data. *Toxicol. Methods* 1:30–43.

Basketter, D. A., Selbie, E., Scholes, E. W., Lees, D., Kimber, I., and Botham, P. A. 1993. Results with OECD recommended positive control sensitizers in the maximization. Buehler and local lymph node assays. *Food Chem. Toxicol.* 31:63–67.

Basketter, D. A., Scholes, E. W., and Kimber, I. 1994. The performance of the local lymph node assay with chemicals identified as contact allergens in the human maximization test. *Food Chem. Toxicol.* 32:543–547.

Basketter, D. A., Gerberick, G. F., Kimber, J., and Loveless, S. E. 1996. The local lymph node assay: a viable alternative to currently accepted skin sensitization tests. *Food Chem. Toxicol.* 34:985–997.

Bekesi, J. G., Holland, J. F., Anderson, H. A., Fischbein, A. S., Rom, W., Wolff, M. S., and Selikoff, I. J. 1978. Lymphocyte function of Michigan dairy farmers exposed to polybrominated biphenyls. *Science* 199:1207–1209.

Bigazzi, P. E. 1988. Autoimmunity induced by chemicals. *J. Toxicol. Clin. Toxicol.* 26:125–156.

Blaikie, L., Morrow, T., Wilson, A. P., Hext, P., Hartrop, P. J., Rattray, N. J., Woodcock, D., and Botham, P. A. 1995. A two-center study for the evaluation and validation of an animal model for the assessment of the potential of small molecular weight chemicals to cause respiratory allergy. *Toxicology* 96:37–50.

Bloom, J. C., Thiem, P. A., and Morgan, D. G. 1987. The role of conventional pathology and toxicology in evaluating the immunotoxic potential of xenobiotics. *Toxicol. Pathol.* 15:283–293.

Botham, P. A., Basketter, D. A., Maurer, T., Mueller, K. D., Potokar, M., and Bontinck, W. J. 1991. Skin sensitization—A critical review of predictive test methods in animals and man. *Food Chem. Toxicol.* 29:275–286.

Boyer, I. J. 1989. Toxicity of dibutyltin, tributyltin, and other organotin compounds to humans and to experimental animals. *Toxicology* 55:253–298.

Buehler, E. V. 1965. Delayed contact hypersensitivity in the guinea pig. *Arch. Dermatol.* 91:171–175.

Buehler, E. V. 1985. A rationale for the selection of occlusion to induce and elicit delayed contact hypersensitivity in the guinea pig. A prospective test. *Curr. Probl. Dermatol.* 14:39–58.

Chin, Y.-H., Falanga, V., Taylor, J. R., Cai, J.-P., and Bax, J. 1990. Adherence of human helper/memory T-cell subsets to psoriatic dermal endothelium. J. Invest. Dermatol. 94:413–417.

Cohen, I. R., and Miller, A. 1994. *Autoimmune disease models: A guidebook.* San Diego: Academic Press.

Condemi, J. J. 1992. The autoimmune diseases. *J. Am. Med. Assoc.* 268:2882–2892.

Dean, J. H., and Murray, M. J. 1991. Toxic responses of the immune system. In *Casarett and Doull's toxicology: The basic science of poisons*, eds. C. D. Klaassen, M. O. Amdur, and J. Doull, 4th ed. pp. 282–333. New York: Macmillan.

Dean, J. H., Cornacoff, J. B., Barbolt, T. A., Gossett, K. A., and LaBrie, T. 1992. Preclinical toxicity of IL-4; A model for studying protein therapeutics. *Int. J. Immunopharmacol.* 14:391–397.

Descotes, J. 1986. Pseudo-allertic drug reactions. *Clin. Res. Pract. Drug Reg. Affairs* 4:75–84.

Descotes, J. 1992. The popliteal lymph node assay; A tool for studying the mechanisms of drug-induced autoimmune disorders. *Toxicol. Lett.* 64/65:101–107.

Du, D. L., Volpe, D. A., Greishaber, C. K., and Murphy, M. J., Jr. 1990. In vitro myelotoxicity of 2',3'-dideoxynucleosides on human hematopoietic progenitor cells. *Exp. Hematol.* 18:832–836.

Fischer, E. I., and Hincks, J. R. 1992. Characterization of the in vitro myelotoxicity in WIN 59075 (SR4233). *Proc. Am. Assoc. Cancer Res.* 33:502.

Fischer, E. I., Hoekstra, S. J., Roberts, A. E., and Hincks, J. R. 1995. In vitro myelotoxicity studies of the thioxanthone anti-cancer agent, WIN 33377, to support a clinical pharmacokinetically guides dose escalation. *Proc. Am. Assoc. Cancer Res.* 36:368.

Gerberick, G. F., House, R. V., Fletcher, E., and Ryan, C. A. 1992. Examination of the local lymph node assay for use in contact sensitization risk assessment. *Fundam. Appl. Toxicol.* 19:438–445.

Gerberick, G. F., Cruse, L. W., Ryan, C. A., Sikorski, E. E., Miller, C. M., and Ridder, G. M. 1996. Differential expression of T cell activation markers (CD62L, CD44, CD45RB, CD25) in draining lymph nodes of mice following exposure to allergens and irritants. *Toxicologist* 30:95.

Harvey, P., Kirshner, J., and Early, A. P. 1984. Leukemia cell cultures. In *Hematopoiesis*, ed. W. Golde, pp. 243–268. New York: Churchill-Livingston.

Hincks, J. R., Adlakha, A., Cook, C. A., Johnson, C. S., Furmanski, P., and Gibson, N. W. 1990. *In vitro* studies on the mechanism of action of hepsulfam in chronic myelogenous leukemia patients. *Cancer Res.* 50:7559–7563.

Holsapple, M. P., Munson, A. E., Munson, J.A., and Bick, P. H. 1983. Suppression of cell-mediated immunocompetence after subchronic exposure to diethylstilbestrol in female B6C3F1 mice. *J. Pharmacol. Exp. Ther.* 227:130–138.

Jalkanen, S. T., and Butcher, E. C. 1985. In vitro analysis of the homing properties of human lymphocytes: Developmental regulation of functional receptors for high endothelial venules. *Blood* 66:577–582.

Kammuller, M. E., Thomas, C., De Bakker, J. M., Bloksma, N., and Seinen, W. 1989. The popliteal lymph node assay in mice to screen for the immune disregulating potential of chemicals—A preliminary study. *Int. J. Immunopharmacol.* 11:293–300.

Karol, M. H., Stadler, J., and Magreni, C. 1985. Immunotoxicologic evaluation of the respiratory system: Animal models for immediate- and delayed-onset pulmonary hypersensitivity. *Fundam. Appl. Toxicol.* 5:459–472.

Kimber, I., and Basketter, D. A. 1992. The murine local lymph node assay: A commentary on collaborative studies and new directions. *Food Chem. Toxicol.* 30:165–169.

Kimber, I., Mitchell, J. A., and Griffin, A. C. 1986. Development of a murine local lymph node assay for the determination of sensitizing potential. *Food Chem. Toxicol.* 24:585–586.

Kimber, I., Hilton, J., and Weisenberger, C. 1989. The murine local lymph node assay for identification of contact allergens: A preliminary evaluation of in situ measurement of lymphocyte proliferation. *Contact Dermatitis* 21:215–220.

Kimber, I., Hilton, J., and Botham, P. A. 1990. Identification of contact allergens using the murine local lymph node assay: Comparisons with the Buehler occluded patch test in guinea pigs. *J. Appl. Toxicol.* 10:173–180.

Kimber, I., Dearman, R. J., Scholes, E. W., and Basketter, D. A. 1994. The local lymph node assay: Developments and applications. *Toxicology* 93:13–31.

Kimber, I., Hilton, J., Dearman, R. J., Gerberick, G. F., Ryan, C. A., Basketter, D. A., Scholes, E. W., Ladics, G. S., Loveless, S. E., and House, R. V. 1995. An international evaluation of the murine local lymph node assay and comparison of modified procedures. *Toxicology* 103:63–73.

Kuper, C. F., Schuurman, H. J., and Vos, J. G. 1995. Pathology in immunotoxicology. In *Methods in immunotoxicology*, eds. G. R. Burleson, J. H. Dean, and A. E. Munson, pp. 397–436. New York: Wiley-Liss.

Lew, F., Tsang, P., Holland, J. F., Warner, N., Selikoff, I. J., and Bekesi, J. G. 1986. High frequency of immune dysfunctions in asbestos workers and in patients with malignant mesothelioma. *J. Clin. Immunoll.* 6:225–233.

Loveless, S. E., Ladics, G. S., Gerberick, G. F., Ryan, C. A., Basketter, D. A., Scholes, E. W., House, R. V., Hilton, J., Dearman, R. J., and Kimber, I. 1996. Further evaluation of the local lymph node assay in the final phase of an international collaborative trial. *Toxicology* 108:141–152.

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

Luster, M. I., Portier, C., Pait, D. G., White, K. L., Gennings, C., Munson, A. E., and Rosenthal, G. J. 1992. Risk assessment in immunotoxicology. I. Sensitivity and predictability of immune tests. *Fundam. Appl. Toxicol.* 18:200–210.

Luster, M. I., Porteir, C., Pait, D. G., Rosenthal, G. J., Germolec, D. R., Corsini, E. Blaylock, B. L., Pollock, P., Kouchi, Y., Craig, W., White, K. L., Munson, A. E., and Comment, C. E. 1993. Risk assessment in immunotoxicology II. Relationship between immune and host resistance tests. *Fundam. Appl. Toxicol.* 21:71–82.

Magnusson, B., and Kligman, A. M. 1969. The identification of contact allergens by animal assay. The guinea pig maximization test. *J. Invest. Dermatol.* 52:268–276.

Montelius, J., Wahlkvist, H., Boman, A., Fernstrom, P., Grabergs, L., and Wahlberg, J. E. 1994. Experience with the murine local lymph node assay: Inability to discriminate between allergens and irritants. *Acta Dermatol. Venereol.* (Stockh.) 74:22–27.

National Academy of Science. 1983. Risk Assessment in the Federal Government: Managing the Process. Washington, DC: National Academy Press.

National Research Council. 1992. *Biologic markers in immunotoxicology*, pp. 33–82. Washington, DC: National Academy Press.

Office of Technology Assessment. 1991. Identifying and Controlling Immunotoxic Substances—Background Paper. OTA-BP-BA-75. Washington, DC: U.S. Government Printing Office.

Pestka, J. J., Tai, J. H., Witt, M. F., Dixon, D. E., and Forsell, J. H. 1987. Suppression of immune response in the B6C3F1 mouse after dietary exposure to the fusarium mycotoxins deoxyvalenol (vomitoxin) and zearanelone. *Food Chem. Toxicol.* 25:297–304.

Rose, N. R., and Mackay, I. R. 1992. The immune response in autoimmunity and autoimmune disease. In *The autoimmune disease II*, eds. N. R. Rose and I. R. Mackay, pp. 1–26. San Diego: Academic Press.

Rosenberg, N. L., Ringel, S. P., and Kotzin, B. L. 1987. Experimental autoimmune myositis in SJL/J mice. *Clin. Exp. Immunol.* 68:117–129.

Schuurman, H. J., Kuper, C. F., and Vos, J. G. 1994. Histopoathology of the immune system as a tool to assess immunotoxicity. *Toxicology* 86:187–212.

Shoenfeld, Y., and Isenberg, D. 1989. *The mosaic of autoimmunity*. Amsterdam: Elsevier.

Sikorski, E. E., Gerberick, G. F., Ryan, C. A., Miller, C. M., and Ridder, G. M. 1996. Phenotypic analysis of lymphocyte subpopulations in lymph nodes draining the ear following exposure to contact allergens and irritants. *Fundam. Appl. Toxicol.* 34:25–35.

Tubbs, R. R., Gephardt, G. N., and Petras, R. E. 1986. Atlas of immunohistology. Chicago: American Society of Clinical Pathology Press.

Urso, P., and Gengozian, N. 1980. Depressed humoral immunity and increased tumor incidence in mice following in utero exposure to benzo(a)pyrene. *J. Toxicol. Environ. Health* 6:569–576.

van der Laan, J. W., and Dean, J. H. 1996. Immunotoxicity of pharmaceuticals: An introduction to the workshop. *Drug Information J.* 30:269–270.

van der Laan, J. W., Van Loveren, H., Vos, J. G., Dean, J. H., and Hastings, K. L. 1997. Immunotoxicity of pharmaceuticals. Current knowledge, testing strategies, risk evaluation and consequences for human health. *Drug Information Journal.* 31:1301–1305.

van Loveren, H., Vos, J. G., and De Waal, E. J. 1996. Testing immunotoxicity of chemicals as a guide for testing approaches for pharmaceuticals. *Drug Information Assoc. J.* 30:275–279.

Verdier, F., Chazal, I., and Descotes, J. 1994. Anaphylaxis models in the guinea pig. *Toxicology* 93:55–61.

White, K. L., Gennings, D., Murray, M. J., and Dean, J. H. 1994. Rat immunotoxicology validation study. *Toxicol. In Vitro.* 8:957–961.

Wolfgram, L. J., Beisel, K. W., and Rose, N. R. (1985. Heart-specific autoantibodies following murine coxsackie virus B3 myocarditis. *J. Exp. Med.* 161:1112–1121.